
The Brunswik Society

Newsletter

Vol. 29, Nov. 2014

To get information about future Newsletters you are welcome to join (free of charge) the email list of the Brunswik Society at:

<http://www.brunswik.org/>

Table of Contents

Foreword

Kaufmann, E., Athanasou, J. A., & Hamm, R. M..... 4

Contributions

Adderley, U., & Thompson, C.
Community Nurses' Judgment and Decision Making for Managing Venous Leg Ulceration:
A Judgment Analysis..... 5

Akhuly, A.
A Competency based Approach to Recruitment Decisions: Through the Axes of the 'How'
and the 'What'..... 8

Araújo, D.
From Representative Design to Affordance-Based Design in Sport..... 11

Bass, E J., Baumgart, L. A., & Klein Shepley, K.
The Effect of Information Analysis Automation Display Content on Human Judgment
Performance in Noisy Environments..... 14

Dhami, M. K.
Berndt Brehmer on Interpersonal Conflict Theory..... 17

Hall, J. A., & Pennington, N.
Using the Lens Model to Assess Personality through Facebook Profile Pages..... 20

Hamm, R. M.
Tool to Produce Desired Cue-Cue Intercorrelations in a Stimulus Set..... 22

Herm, S.
Information Aggregation in Online Communities..... 25

Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D.
A Lens Model Approach to Accurate Personality Judgments: The Case of Neuroticism and
Situational Relevance..... 27

Huttin, C.
Judgment Studies and Economics of Information Exchanges in Healthcare..... 30

Madison, G.
Testing an Evolutionary Theory of Human Rhythm and Groove: Tapping Musicians'
Implicit Knowledge through Egon Brunswik's Lens Model..... 31

Mumpower, J. L.
News from the Past Year..... 36

Nash, U. W.
Probabilistic Functionalism and Systematic Error in the Wisdom of Crowds..... 37

Nordvall, A.-C.
Recycling is Good!..... 41

Orquin, J. L.
How do We Judge the Healthiness of Food?..... 43

Reinsch, N. L., & Gardner, J. A.
Do Communication Abilities Affect Promotion Decisions? Some Data from the C-Suite..... 45

Sjödahl, L.
The Decision Maker as a Planner..... 46

Wieser, M.
The Structure of Brunswik’s “Lens”: A Media-Historical Analysis..... 52

Wigton, B., & Tape, T.
News from Nebraska..... 55

Zlatkin-Troitschanskaia, O., Kuhn, C., & Toepper, M.
The German Research Program KoKoHs – Theoretical Concept, Assessment Design, and Validation Approach in Modeling and Measuring Competencies in Higher Education..... 56

Imprint

The Brunswik Society Newsletter, ISSN 2296-9926
is published yearly
Editorial staff:
Kaufmann, E., Athanasou, J. A., & Hamm, R. M.

Contact:

The Brunswik Society Newsletter
c/o University of Zurich
Institute of Education, University of Zurich,
Kantonsschulstr. 3
CH-8006 Zurich
Switzerland

Foreword

It is a great pleasure to present again a rich Brunswik Society Newsletter. This issue covers subjects from the past, going back to the history of the lens model (see Wieser), through to modern life (see Hall & Pennington).

Unfortunately, this year the Brunswik expert Berndt Brehmer passed away. In his honor, several contributions refer to his work. Dhami's contribution presents Brehmer's research on Interpersonal Conflict Theory. Araújo founded his research project on Brehmer's work. In Hamm's contribution you will find a tool for controlling cue-cue intercorrelations or simply for doing research in line with Brehmer. We hope that reading this issue of the Brunswik Newsletter will help researchers learn more about Brunswik's theory and become inspired by Brehmer's wonderful legacy.

Beside these contributions, this year a contribution (see Zlatkin-Troitschannskaia and team) is included that looks for collaboration in relation to Brunswik knowledge – indicating the relevance and interest in Brunswik research.

We hope that the richness of the contributions included in this year's Newsletter inspires further thoughts, links and discussions for further research.

To keep-up with the Brunswik Society, please follow the Brunswik Society mailing list, which you will find at: www.brunswik.org - we are looking forward to hearing your thoughts, ideas, etc., using this list.

Many thanks to all authors for their contributions!

Sincerely,

Esther Kaufmann, James A. Athanasou and Robert M. Hamm

Thank you to Tom Stewart, the webmaster of the Brunswik Society, for providing web access to the Newsletter.

Contributions

Community Nurses' Judgment and Decision Making for Managing Venous Leg Ulceration: A Judgment Analysis

Una Adderely

*School of Healthcare, University of Leeds,
UK*

Carl Thompson

*Department of Health Sciences, York University,
UK*

Contact: u.adderley@leeds.ac.uk / Carl.thompson@york.ac.uk

We explored how community nurses manage the uncertainty associated with diagnosing and treating venous leg ulceration. Despite being an important part of community nurses' workload (Posnett & Franks, 2008), quality of diagnosis and treatment is variable and often below that which should be expected (Royal College of Nursing, 2008; Srinivasaiah et al., 2007; Vowden & Vowden, 2009).

We took a Brunswikian approach because clinical judgment occurs in uncertain environments which bear little resemblance to controlled experimental settings. We sought to understand the relationship between the nurses, the cues in the clinical environment, and their judgment accuracy (Goldstein, 2007). Leg ulcer management involves tasks such as judging whether to carry out certain tests, interpreting the results, how to present information to the patient, and when to commence therapy. Accuracy was likely to depend on a nurse's ability to prioritise relevant information and disregard the irrelevant (Cooksey, 1996).

We used Judgment Analysis to explore the performance of 36 community nurses. In order to achieve a representative sample of judges (Cooksey, 1996) we recruited 18 tissue viability specialist nurses and 18 generalist community nurses who made these sorts of judgments in real life and thus had a high level of familiarity with the task. Task congruence was assured by creating a representative sample of the clinical environment. This was using 110 clinical scenarios generated from real patient cases. The proportions of different types of leg ulcer mirrored the UK leg ulcer population and each written scenario was accompanied by a photograph of the wound to add visual information.

The nurses viewed the scenarios online and made a diagnosis and a treatment judgment for each scenario. These judgments provided the data for the subject side of a double system lens model (Cooksey, 1996). A nominal group (Black, 2006) consensus panel was formed which consisted of four community tissue viability nurses

with advanced relevant clinical and research experience. These nurses were asked to independently complete the judgment task before they met for the consensus panel. At the consensus panel meeting they were presented with each scenario, informed of their range of individual answers and asked to agree a group answer. Any disagreements were resolved by discussion. These consensus judgments were used to provide the 'ecology' data for the left side of the lens model.

Logistic regression models were constructed to examine nurses' use of the information in the scenarios (Cooksey, 1996; Stewart, 2004). Differences between generalist and specialist tissue viability nurses and between nurses with different levels of education were explored using paired t-testing and ANOVA (Field, 2005).

R_e for diagnosis was 0.63, indicating that the nurses could be reasonably expected to have an R_a of up to 0.63. The nurses achieved an overall accuracy of 0.48. For the treatment judgment, the predictability was 0.88 so there was less uncertainty, but the nurses achieved an accuracy of only 0.49. For both judgments, the specialist nurses were more accurate than the generalist nurses. Level of education was not a predictor of better accuracy.

The nurses gave the appropriate weight to the most important diagnostic cue but under-weighted other important cues and over-weighted less important cues. For the treatment judgment, the nurses gave insufficient weight to the most important cues but over-weighted less important cues.

In this study the judgments of both generalist and specialist community nurses were not as accurate as they could be. It was particularly surprising that treatment judgments were no more accurate than diagnosis judgments, despite there being less uncertainty in the treatment clinical environment. This might be related to the cue weights but this does not explain why specialist nurses were more accurate than generalist nurses.

Our study is innovative in being the first judgment analysis study to focus on tissue viability, community nurses and the impact of expertise on the management of leg ulceration. Our study is also the first judgment analysis study in nursing to use photography to enhance representativeness. The use of computerised simulation enabled the presentation of a larger number of scenarios than standard methods would have permitted, thus we were able to generate more stable logistic regression estimates. We do not know whether digital presentation impacted on participants' performance. Future Judgment Analysis studies might like to consider the tradeoff between presenting sufficient scenarios to achieve stable logistic regression estimates and the demands on the participants associated with the manner in which the scenarios are presented. We also used a much larger number of cues than the 7 (+) cues recommended for Judgment Analysis research (Cooksey, 1996; Miller, 1956). It has been suggested that even when many cues are available, participants typically use fewer than 10 cues (Roose & Doherty, 1976). The results of this thesis support these findings.

This study exposed the complexity of the clinical environment surrounding the management of leg ulceration and to set out models for diagnostic judgment and treatment choices for venous leg ulceration. These models provide a starting point for developing robust strategies for supporting community nurses' judgment and decision

making. Such strategies will require investigation to assess their potential usefulness but they offer the possibility of more clinically and cost effective care for patients with venous leg ulcers

References:

- Black, N. (2006). Consensus development methods. In C. Popp & N. Mays (Eds.), *Qualitative Research in Health Care*. 3ed. Oxford: Blackwell Publishing.
- Cooksey, R. W. (1996). *Judgment Analysis: Theory, Method and Applications*. New York, Academic Press.
- Field, A. (2005). *Exploring Data. Discovering Statistics using SPSS*. London: Sage Publications.
- Goldstein, W. M. (2007). Social Judgment Theory: Applying and extending Brunswik's Probabilistic Functionalism. In D. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 37-67). Oxford: Blackwell.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63, 81-97.
- Posnett, J., & Franks, P. (2008). The burden of chronic wounds in the UK. *Nursing Times*, 104, 44-45.
- Roose, J. E., & Doherty, M. E. (1976). Judgment theory applied to the selection of life insurance salesman. *Organizational Behaviour and Human Performance*, 16, 231-249.
- Royal College of Nursing (2008). *National Audit: The management of venous leg ulcers*. www.rcn-audit.org.uk.
- Srinivasaiah, N., Dugdall, H., Barrett, S., & Drew, P. J. (2007). A point prevalence survey of wounds in north-east England. *Journal of Wound Care*, 16, 413-419.
- Stewart, T. (2004). *Notes on a form of the lens model equation for logistic regression analysis*. The Brunswik Society Meeting.
- Vowden, K., & Vowden, P. (2009). The prevalence, management and outcome for patients with lower limb ulceration identified in a wound care survey within one English health care district. *Journal of Tissue Viability*, 18, 13-19.

**A Competency based Approach to Recruitment Decisions:
Through the Axes of the ‘How’ and the ‘What’**

Ajanta Akhuly

*Indian Institute of Technology, Bombay,
India*

Contact: ajanta.iitb@gmail.com

My interest in Brunswik’s Lens model was elicited in my doctoral dissertation which was titled “A Competency Based Approach to Recruitment Decisions: From the Lens Model to the Practice Turn” (Akhuly, 2013). The dominant interest of the thesis has been to understand how recruiters arrive at hiring decisions. “Decision making” was argued to be the most important phenomenon in personnel selection because recruitment is essentially a decision making process. Thus, this work found itself at the interface of Human Resource Management and decision making.

For the sake of deploying a Brunswikian lens model, a competency based approach was adopted to generate the predictors on the basis of which candidates are chosen. Qualitative interviews were conducted to identify competencies that recruiters look for while hiring Human Resource (HR) professionals. An analysis of the interviews showed that HR incumbents should possess three major competencies (people, individual and business orientation) with background being an important indicator in the Indian context. Based on the qualitative study, a field experiment was designed on the basis of the Lens model (for full paper, see Akhuly & Gupta, 2014). Vignettes were designed where each vignette is a combination of all the 4 competencies identified varied at a “high” or “low” level. These 16 vignettes were administered to another set of thirty five HR professionals. A five point rating scale was used to report the certainty of recruiting the hypothetical candidates on each vignette. After they arrived at a decision for all the 16 vignettes, they had to rank order a set of competencies they thought important.

The results of ANOVA (people, individual & business have significant main effect and interaction effects) explicated that possessing some competencies increases an HR’s certainty of getting recruited. Cluster analysis was done to find out judges who have similar decision making patterns. In the overall sample four clusters were formed. Estimates of model-fit of the binary logistic regression carried out on each cluster showed how each cluster was in fact different from the other clusters. As a measurement of accuracy, cues that were given importance during the judgment task (weights obtained in logistic regression) were compared with cues that were thought to be important (weights obtained from ranking competencies). Results taken together show that the way interviewers actually evaluate candidates may differ from the way in which they believe they evaluate. The coefficient of inter-rater reliability indicated “poor” extent of agreement among judges. Most of the estimates calculated in this study (for example, mapping what they think with what they do; inter-rater agreement) adhere to coherence as a criterion of decision making. Results of this study were interpreted in light of Brunswik’s Lens model.

In some sense, in order to push my boundaries of thinking, the aspiration of the dissertation has been to understand the “nature” of decision making. An attempt to elucidate what might be called the ‘inherent properties’ of decision making. In order to grapple with this, I tried to delineate what the lens model does in terms of its methodology, in terms of its philosophy (the strands of decision making that it pulls out) and in terms of its operationalization. The anxiety about which strands of decision making the lens model can throw light on is reflected by Brehmer and Brehmer’s observation, “It is far from clear what kinds of questions have been asked about the nature of human judgment in studies of policy capturing and what kinds of questions can be answered from the existing results. Indeed, it is not even clear that all (or even a majority) of the studies actually ask any fundamental questions at all about the nature of human judgment ...” (1988 cited in Cooksey, 1996, p. 58).

If I attempt to highlight some of the properties of decision making that the lens model reflects which I have encountered in my study. First and foremost being the Lens model’s essential premise of the linearity assumption where both environments and judges are often well modeled by linear functions (Karelaia & Hogarth, 2008). While Gigerenzer and Kurz (2001) argue that linear models do not, per se, specify the cognitive processes, Katsikopoulos (2009) counter-argues that, “weighting and summing cues can be seen as processes... cues are first weighted in decreasing order of some subjective measure of importance, and then cues are summed in decreasing order of weight”. The second important property is Brunswik’s emphasis on the importance of “perception” in judgment (Harvey, 2001). While, the role of perception lies in the judge’s ability to infer the unseen from the “seen”, thinking on the other hand, was described as certainty-gearred, machine-like. Decision making was one domain of reasoning that Brunswik saw as cutting across the dichotomy of perception and thinking (Doherty & Kurz, 1996). Finally, if the quintessence of decision making is about choice, then we always want to know whether that choice was right or wrong, i.e., whether the judgment itself was right or wrong. Hammond’s claim is that conclusions about the competence of judgments and decisions will depend upon the selection of coherence or correspondence as the criterion (Hammond, 2008 cited in Dunwoody, 2009). Researchers in the Brunswikian tradition tend to emphasize the correspondence of judgment with ecological criteria (Hammond, 2007 cited in Dunwoody, 2009).

Though the points that I have summarized are in no sense exhaustive, and my comprehension of the Lens model is way too inadequate. However, I was still invested in understanding the “nature of decision making”. During that time I had stumbled upon W. T. Stace’s (1920) “A Critical History of Greek Philosophy”. Stace (1920, p. 274) elaborates that Matter and Form are the fundamental categories of Aristotle’s philosophy. First of all, matter and form are inseparable. We think of them as separate in order to understand them clearly (Stace, 1920, p. 275). The form includes all the qualities of the thing. The matter is what has the qualities. For the qualities are all universals. A thing without qualities cannot exist, nor qualities without a thing. And this is the same as saying that form and matter cannot exist separately (Stace, 1920, p. 278). For example, we know that there are no such things as squares, circles, and triangles. There are only square objects, circular objects, etc. Geometry is quite right to treat shapes as if they existed by themselves, but it is nevertheless dealing with mere abstractions (Stace, 1920, p. 275). For the purpose of the dissertation, I had attempted to think about the phenomenon of “decision making” through two axes: the

“what” and the “how” question. I have argued that the “what” and the “how” are intricately intertwined like “matter” and “form”.

I am not sure if I am drawing the analogy correctly from Aristotle, but if we think of “recruitment” as the “matter” and “decision making” as the process / “form”. Since “decision making” is a process it needs a content/context on which it has to operate. The context of “decision making” can be anything (personnel selection, consumer decision making, marriage). But, whatever may be the context one cannot talk about the “process” of decision making without talking about the context for which the decision is made. It’s a bit like “matter” (recruitment scenario) and “form” (the process of making decisions) where one cannot exist without the other. Another way to think about this project is in terms of the “what” and the “how” question. Very simply put, if I ask the recruiters “what” they look for in candidates while hiring? (they immediately start talking in terms of competencies). But when I ask “how” do you know, they start elaborating the process. The “what” is the content, without the “what” the “how” is just not possible. However, the “what” and the “how” question is distinctly different.

Within the lens model paradigm there is definitely a certain emphasis on “what” constitutes the “true” state (in our case the person to be recruited). As documented in Maule (2001), Dhami and Harries argue that, “any situation, object, or person that is to be judged can be represented in terms of a number of different cues, each varying in terms of the extent to which they reflect the “true” state. Judgment involves identifying and integrating this information, weighting and evaluating one or more cues to arrive at an overall evaluation”. Most often we want to know “what” actually matters. In our study, the “what” question was answered by decomposing the judgments through methods such as ANOVA, cluster analysis and binary logistic regression to find out which competencies really matter while hiring.

I have a lingering feeling that the strands that constitute the “nature” of decision making within the lens model paradigm, has revolved more around the “what” rather than the “how” question. I think that if we deal with these two questions separately (the “what” and the “how”), we may be able to dissect the confusions that have risen related to the “nature” of decision making. Though these two questions look incredibly simple and very obvious, however, there might be value in thinking about them separately.

References:

- Akhuly, A. (2013). *A competency based approach to recruitment decisions: From the Lens Model to the practice turn* (Unpublished doctoral dissertation). Indian Institute of Technology, Bombay, India.
- Akhuly, A., & Gupta, M. (November, 2014). Competency based recruitment decisions: A Lens Model Approach. *The Psychologist-Manager Journal*, 17(4). [paper in press]
- Cooksey, R. W. (1996). *Judgment analysis: Theory, methods, and applications*. UK: Emerald.
- Doherty, M. E., & Kurz, E. M. (1996). Social judgment theory. *Thinking & Reasoning*, 2(2-3), 109-140.
- Dunwoody, P. T. (2009). Theories of truth as assessment criteria in judgment and decision making. *Judgment and Decision Making*, 4(2), 116-125.
- Gigerenzer, G., & Kurz, E.M. (2001). Vicarious functioning reconsidered: A fast and frugal lens model. In K. R. Hammond & T. Stewart (Eds.), *The essential Brunswik: Beginning, explications, applications* (pp. 342-347). Oxford, UK: Oxford University Press.
- Harvey, N. (2001). Studying judgment: General issues. *Thinking & Reasoning*, 7(1), 103-118.
- Karelaia, N., & Hogarth, R. M. (2008). *Determinants of linear judgment: A meta-analysis of Lens Model studies*. *Psychological Bulletin*, 134(3), 404-426.
- Katsikopoulos, K. V. (2009). The conceptual connection between lens models and fast and frugal heuristics: A process approach. *Theory and Psychology*, 19(5), 688-697.

Maule, J. A. (2001). Studying judgment: Some comments and suggestions for future research, *Thinking & Reasoning*, 7(1), 91-102.

Stace, W. T. (1920). *A critical history of Greek philosophy*. Glasgow: The University Press.

From Representative Design to Affordance-Based Design in Sport

Duarte Araújo

*CIPER, Faculdade de Motricidade Humana, Spertlab, Universidade de Lisboa,
Portugal*

Contact: daraujo@fmh.ulisboa.pt

This contribution for the Brunswik Society Newsletter is a re-arrangement of certain parts of the article:

Araújo, D. and Davids, K. (in press). Towards a theoretically-driven model of correspondence between behaviours in one context to another: Implications for studying sport performance. *International Journal of Sport Psychology*.

Nearly a decade ago we published an experiment on decision making in the sport of sailing, where the task for participants was to perform in a computerised, simulated regatta (Araújo, Davids, & Serpa, 2005). The conceptual foundations of the study were based on Brehmer's work (e.g., Brehmer, 1996), which has demonstrated the external validity of performing on computer simulations, where external validity is the degree to which the results can be generalized to other samples and situations. In our case, external validity was obtained with a high correlation between the level of expertise of the sailors and their performance on the sailing simulator.

However, the measurement of "Brunswik's" representativeness can be very difficult in many situations, such as those of direct competition with an opponent or opposing team in sport (e.g., Araújo et al., 2007). For example, how is it possible to know in advance the composition of an adversary team? Sport analysts may speculate, but they cannot calculate the precise ecological validity of cues, their interrelation, and the overall uncertainty of a future competitive game, without knowing which cues will be perceived by performers.

Another point that demands an approach that goes beyond Brunswik's (1956) "representative design" is that by randomly sampling cues in an environment, the researcher may be excluding those unique cues that are available for attuned, expert participants. This means that, more than sampling cues in an environment, it may be more relevant to select affordances (i.e., possibilities for action). For Brunswik, how the individual acts was not very relevant. Brunswik, as well as his influential co-author Edward C. Tolman (e.g., Tolman & Brunswik, 1935), clearly rejected the so-called molecular approaches and the particular behaviours of organisms were not a key issue. What mattered for them were the consequences of the behavior (the molar level), i.e., what the organisms were getting at, not how they were getting there (see

also Goldstein, 2004, p. 40). For sport scientists, as well as for Gibsonian ecological psychologists, the problem of “how to act” in a given situation is a major concern, given that the structure of action is an expression of goal-directness.

The representative task as a “modified environment” implies a specific adaptation of the athlete, since both contexts (the representative task and the performance context) may have different affordances (Araújo et al., 2007). This implies a theoretical change, from understanding an environment that interacts with a performer (i.e., two separate systems that interact) to a more holistic view, which understands performer and environment as a single coherent system to be studied.

Finally, there are contexts which are not representative, or even which may never be observed in a competition (the context towards which one wants to generalize), but which can contribute to understanding and testing the performance of athletes in competition. For example, in team sports like futsal or basketball, coaches in some specific circumstances prepare their five-player teams with overload training against teams with 6 and 7 players. In competition this will never occur, but coaches use this method to train their players to work harder to create space away from opposing players. The overloading of opposing players is undertaken in order for trainees to become better perceptually attuned to affordances that are relevant in competitive performance environments. We may think about similar non-representative contexts built on virtual environments, where researchers (and coaches) may manipulate circumstances relevant to understand performance in a competition, but that could not be manipulated out of the virtual context.

These limitations of Brunswik’s concept of representative design led us to propose an affordance-based design. In particular, we seek to propose principles to demonstrate the correspondence between behavior in an experimental setting and behavior in a context towards which we want to generalize a behavior measured in an experimental setting.

The interrelated criteria to test behavioral correspondence between contexts are:

1) *Selection of functional (i.e. relevant) affordances.* Even though affordances can be empirically tested (see Fajen, Riley, & Turvey, 2009), their selection should be theoretically driven, even if the researcher is not adopting an ecological psychology standpoint.

2) *Action fidelity.* For Stoffregen et al. (2003) the key aspects of this concept are that: i) perception is defined with respect to behavior, and ii), action fidelity does not mandate a concentration on “stimulus fidelity”, since the environment is defined in behavioral terms (i.e., affordances). Therefore, action fidelity concerns the degree to which actions performed in the experimental setting are related to the actions performed during competition.

3) *Performance achievement.* Achievement is the degree of success obtained when performing a task for a specific goal. It can be seen as a concept similar to ecological rationality. For Gigerenzer, behavior is successful if it is adapted to the structure of the information in the environment in which it is realized (Gigerenzer, Todd & ABC research group, 1999).

In conclusion, Brunswik developed important theoretical principles, which can be integrated into a broader ecological dynamics view enhancing our understanding about how behavior in one context may correspond to behavior in another context. The suggestion is that modeling the “behavioral correspondence between contexts” ensures that such generalizations can be achieved, based on adherence to the intertwined criteria of selection of affordances, action fidelity and performance achievement.

References:

- Araújo, D., Davids, K., & Passos, P. (2007). Ecological validity, representative design and correspondence between experimental task constraints and behavioral settings. *Ecological Psychology, 19*, 69-78.
- Araújo, D., Davids, K., & Serpa, S. (2005). An ecological approach to expertise effects in decision-making in a simulated sailing regatta. *Psychology of Sport and Exercise, 6*, 671-692.
- Brehmer, B. (1996). Man as a stabiliser of systems: From static snapshots of judgment processes to dynamic decision making. *Thinking and Reasoning, 2*, 225-238.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley: University of California Press.
- Fajen, B., Riley, M., & Turvey, M. (2009). Information, affordances and the control of action in sport. *International Journal of Sport Psychology, 40*, 79-107.
- Goldstein, W. (2004). Social Judgment Theory: Applying and extending Brunswik's probabilistic functionalism. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 37-61). Malden: Blackwell publishing.
- Gigerenzer, G., Todd, P. M., & The-ABC-Research-Group (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Stoffregen, T. A., Bardy, B. G., Smart, L. J., & Pagulayan, R. J. (2003). On the nature and evaluation of fidelity in virtual environments. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual and adaptive environments: Applications, implications, and human performance issues* (pp. 111-128). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Tolman, E. C., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review, 42*, 43-77.

**The Effect of Information Analysis Automation Display Content
on Human Judgment Performance in Noisy Environments**

Ellen J. Bass
Drexel University,
USA

Leigh A. Baumgart
NorthShore University HealthSystem,
USA

Kathryn Klein Shepley
The MITRE Corporation,
USA

Contact: Ellen.J.Bass@drexel.edu

This article summarizes Bass et al. (2013) on the use of the Human-Automated Judge Learning (HAJL) (Bass & Pritchett, 2008) to analyze the effect of display interventions with information analysis (IA) automation. Using a triple system design, HAJL models interaction among the environmental criterion, the person's judgment and IA automation's judgment. In the training phase, the automation and the human independently make a judgment. In the interactive learning (IL) phase, the correspondence between their initial judgments measures conflict. After viewing the automation's judgment, the person provides a revised joint judgment. With achievement and conflict, adaptation (correlation between the joint and automation's judgments) and compromise (correlation between initial and joint judgments) measure performance. In the prediction phase, the person makes a judgment independent of the automation and also predicts the automation's judgment. The person's predictive accuracy (correspondence between automation and the prediction), assumed similarity (correspondence between the person's judgment and prediction of the automation), and actual similarity (correspondence between the person's and the automation's judgments) measure performance.

This work investigated the contribution of providing IA automation strategy information, task environment information, or both, on human judgment performance when noisy sensor data are available.

Methods

A Traffic Conflict Prediction System (TCPS) calculated probability of air traffic conflict (PC) judgments using noisy data (Bass & Pritchett, 2008). The environmental criterion (the actual PC) was calculated using the TCPS with no noise. The display content conditions in the IL phase were:

- automation judgment (O) - TCPS's PC
- O plus OE (task environment information related to sensor noise)

- plus OA (TCPS's judgment strategy via intermediate cues (projected positions of ownship and traffic at the point of closest approach, representations of uncertainty, and the ownship protected zone))
- O, OE and OA

180 trials were grouped into 6 sessions of high environmental predictability (0.971, 0.971, 0.969, 0.985, 0.989, and 0.974). In the training and IL phases, participants experienced 6 sessions and 3 in the prediction phase.

A repeated measures, mixed model design was used. Display content condition, session, and the session-display content interaction were fixed effects. Participants were nested within display content and treated as a random effect. Post hoc analysis was conducted using Tukey's Honestly Significant Difference. Wilcoxon Signed Rank tests compared achievement data. The inverse sine transformation was applied to judgments and correlations were transformed using Fisher's r to z_r transformation.

Results

Thirty-two participants made horizontal PC judgments.

In the IL phase, the participants' unaided judgment achievement, r_{a1} ($\mu = 0.49$, $\sigma = 0.18$), was higher than in the training phase ($V = 12652$, $p < 0.001$) likely due to additional practice and from learning from the automation. Average joint judgment achievement, r_{a2} ($\mu = 0.91$, $\sigma = 0.06$) was higher than the average unaided judgment achievement ($V = 0$, $p < 0.001$) and closer to the automation's achievement ($\mu = 0.94$, $\sigma = 0.02$). Levene's test showed the variance of joint judgments was smaller compared to the variance of unaided judgments ($F_{191, 191} = 10.251$, $p < 0.001$) and the floor of judgment achievement was raised from -0.16 (unaided) to 0.66 (joint) across all participants. Display content condition ($F_{3, 168} = 5.979$, $p = 0.001$) and session ($F_{5, 168} = 9.5782$, $p < 0.001$) both impacted joint judgments. Participants needed a session before knowing how to use the output from the TCPS.

The O condition had lower joint achievement than OE ($p = 0.003$) and OEA ($p = 0.02$). There was a trend for OA to be lower than OEA ($p = 0.08$). Automation strategy information did not significantly help compared to when participants were also provided with environment information.

Neither compromise ($\mu = 0.49$, $\sigma = 0.18$) nor adaptation ($\mu = 0.93$, $\sigma = 0.05$) were found to be significantly different across display content condition in the IL phase. Compromise was low and adaptation was high for all (participants were adapting their unaided probability judgments to match the automation's judgments).

Session was significant for compromise ($F_{5, 168} = 2.6983$, $p = 0.023$) and adaptation ($F_{5, 168} = 7.7129$, $p < 0.001$). Session 4 (conducted the day after session 3) was lower than session 1 for compromise ($p = 0.045$). For adaptation, session 1 was lower than session 3 ($p < 0.001$) and session 5 ($p < 0.001$), indicating that as participants gained experience with the automation, their joint judgments more closely corresponded with the automation's.

The participants' unaided judgment achievement in the prediction phase, r_{a1} ($\mu = 0.58$, $\sigma = 0.19$), was not significantly different based on session or display content

condition. Average unaided judgment was higher than during the IL phases, as participants may have learned from the automation. The participants' unaided judgment achievement in the prediction phase, r_{a1} , was higher than their predictive accuracy ($\mu = 0.54$, $\sigma = 0.18$) ($V = 0$, $p < 0.001$). Predictive accuracy was not significantly affected by either display content condition or session order. Assumed similarity ($\mu = 0.90$, $\sigma = 0.16$) was higher than actual similarity ($\mu = 0.56$, $\sigma = 0.18$) ($V = 4592$, $p < 0.001$) across all display content conditions. Participants were poorly calibrated with respect to understanding the difference between their performance and the automation's performance. Neither similarity measure was significantly affected by either display content condition or session order.

Discussion

This research sought to investigate the impact of IA automation display content on human judgment performance. Participants provided with environment information had significantly higher joint judgment achievement compared to those provided with only the automation's judgment. Providing automation strategy information in the IL phase did not appear to help participants understand the automation. Participants assumed that their predictions of the automation would be closer than they actually were. These results have implications for IA automation design. Here the automation's imperfect judgments were tied to the noisy input data. Although providing participants with information regarding the automation's judgment strategy may improve performance compared to those receiving only the automation's judgment, this information does not appear as beneficial to participants as information pertaining to the uncertainty in the environment when making their joint judgments. It is possible that additional environment information allowed the participants to exploit the automation more effectively as they understood how automation judgment achievement varied based on factors in the environment.

References:

- Bass, E. J., Baumgart, L. A., & Shepley, K. K. (2013). The effect of information analysis automation display content on human judgment performance in noisy environments. *Journal of Cognitive Engineering and Decision Making*, 7(1), 49-65.
- Bass, E. J., & Pritchett, A. R. (2008). Human-automated judge learning: A methodology for examining human interaction with information analysis automation. *IEEE Transactions on Systems, Man, and Cybernetics*, 38(4), 759-776.

Berndt Brehmer on Interpersonal Conflict Theory

Mandeep K. Dhami
Middlesex University,
UK

Contact: m.dhami@mdx.ac.uk

Over the years, I have found Berndt Brehmer's work fundamental to my understanding of neo-Brunswikian concepts. Brehmer made significant contributions to the introduction of Brunswikian ideas into modern day psychology. In particular, he was a co-author along with Hammond, Stewart and Steinmann on the 1975 book chapter that describes the tenets of *Social Judgment Theory* (SJT). He followed this up in 1988 with a book edited together with Joyce, in which SJT theory and methods were reviewed and applications of SJT were presented. Brehmer also critically commented on Hammond's proposal for '*formal situational sampling*' as an alternative to Brunswik's vision of representative design using substantive situational sampling; describing Hammond's alternative as "no easy road to success" (Brehmer, 1979, p. 198). In addition, Brehmer pointed to the difficulties of learning from outcome feedback alone and highlighted the benefits of *cognitive feedback* (Brehmer, 1980). But, perhaps Brehmer's most significant, and less well-known contribution, was the development and testing of *Inter-personal Conflict Theory* (ICP). Below, I summarize this using excerpts taken from Dhami, M. K., & Olsson, H. (2008). Evolution of the interpersonal conflict paradigm, *Judgment and Decision Making*, 3, 547-569. I hope the next generation of decision researchers will come to learn from and appreciate Brehmer's work, as I have done.

An early review of research using IPC theory published in *Psychological Bulletin* by Brehmer (1976) highlighted the potential of [...] advancing our understanding of cognitive conflict in both laboratory and real world settings.

[...] Brehmer and colleagues conducted an intensive series of studies. In 1976 Brehmer reviewed the research that had been conducted on cognitive conflict using IPC theory. By then, research had examined issues concerning: (a) the structure of cognitive conflicts; (b) the relative importance of the task and the other party in affecting policy change and conflict resolution; (c) the effect of task characteristics on cognitive conflict; and (d) the effect of person characteristics on cognitive conflict. Research had also begun to study (e) how cognitive conflict could be resolved via supports/aids. We describe the main findings below.

First, conflict may persist due to non-systematic cognitive differences even when parties are motivated to agree, and actually do agree in principle. Indeed, while parties reduce the systematic differences in their policies (i.e., there is policy similarity), over time the inconsistency of their policies increases thus leading to little reduction in the amount of conflict although the structure of the conflict has altered (e.g., Brehmer, 1969). This is because parties tend to decrease their dependency on

their old policies at a faster rate than they increase their application of a new policy that is compatible with each other's (e.g., Brehmer, 1972).

Second, policy change itself does not signify willingness to compromise but rather a desire to achieve, although compromise is sought when accuracy is not clearly observable/obtainable. When one party is initially trained in the optimal policy and the other is not, the latter will learn from the former if the task is highly predictable (e.g., Brehmer, 1973a). However, if task predictability is low, the parties start off by decreasing dependency on their initial policies. Here, based on feedback, the party with the optimal policy soon appropriately switches back to his/her original policy, and the other party also learns from feedback (e.g., Brehmer, 1974). When there is no feedback, parties may compromise: this reduces conflict without leading to observable inaccuracy (e.g., Brehmer, 1971).

Third, formal (surface and system) task characteristics can influence each party's policy development and the ease with which they can achieve, and such characteristics alone can explain cognitive conflict. Hammond and Brehmer (1973) did not find much evidence for substantive or content task characteristics influencing cognitive conflict. Surface characteristics refer to the number of cues, the metric level of cues, and the inter-cue correlations, while system characteristics refer to the distribution of cue validities, forms of functions relating cues to the criterion, organizing principles, and task predictability. For example, there is greater agreement despite less reduction of policy differences when the cues are inter-correlated than when they are orthogonal (Brehmer, 1975). This may be because cue inter-correlations enable the parties to achieve with little change of their original policies (Mumpower & Hammond, 1974). In addition, there is less agreement between parties when task predictability is low because each party's policies are less consistent rather because of any systematic differences in their policies (e.g., Brehmer, 1975). Similar findings have been observed for tasks that require policies with nonlinear function forms which tend to be more difficult to develop (e.g., Brehmer, 1973b).

Fourth, traditional individual difference variables such as gender do not affect measures of cognitive conflict (Hammond & Brehmer, 1973).

Finally, cognitive aids may be useful for reducing conflict. Hammond and Brehmer (1973) applied the technique of cognitive feedback (Todd & Hammond, 1965) and developed a cognitive aid to conflict resolution called POLICY.6 This interactive computer program enables parties to express their policies, compare them, change them, and discover the effects of such changes on conflict (see Rohrbaugh, 1988, for group decision support systems). Cognitive feedback involves providing information about the task (i.e., ecological validities, intercue correlations, predictability, and cue-criterion function forms), the party's judgment policy (i.e., utilization validities, cognitive control/consistency, and cue judgment function forms), and the match between them (i.e., achievement, and its linear and nonlinear components) (Balzer, Doherty, & O'Connor, 1989; Doherty & Balzer, 1988). Such feedback can help to speed conflict reduction (Balke, Hammond, & Meyer, 1973).

[...] In 1969, Leon Rappoport warned [...] that "if the cognitive conflict model is to serve as anything more than a laboratory analogue, it must be determined whether socially induced (i.e., "natural") cognitive differences generate the same conflict phenomena as laboratory induced (i.e., "artificial") cognitive differences" (p. 143). In fact, as Brehmer

(1976) noted, many of the findings that were observed in the laboratory on simulated tasks were also obtained in naturalistic environments or real tasks, particularly for use in policy development (e.g., Adelman, Stewart, & Hammond, 1975; Balke et al., 1973; Brown & Hammond, 1968; Steinmann, Smith, Jurdem, & Hammond, 1975). Brehmer (1976) concluded his review with avenues for future research including examining the antecedents and consequences of policy inconsistency, and further analysis of real world conflicts.

Reference:

Excerpts taken from: Dhami, M. K., & Olsson, H. (2008). Evolution of the interpersonal conflict paradigm. *Judgment and Decision Making*, 3, 547-569.

**Using the Lens Model to Assess Personality
through Facebook Profiles Pages**

Jeffrey A. Hall & Natalie Pennington
*Communication Studies at the University of Kansas,
USA*

Contact: hallj@ku.edu; natpenn@ku.edu

When we embarked on our Facebook Lens Model investigation in Spring 2011, we were hoping to be the first research team to document the role of status updates on impression management and impression formation. Drawing from Warranting Theory (Walther & Parks, 2002), we hoped to identify cues with high warranting value, such as the number of Facebook friends for judging extraversion (Tong et al., 2008), that were also diagnostic warrants, or cues linking user personality with stranger perception.

Our research (Hall, Pennington, & Lueders, 2014; Hall & Pennington, 2012) extended previous work by exploring seven traits (i.e., openness, conscientiousness, extraversion, agreeableness, neuroticism, humor, narcissism) and 53 cues on a profile page. We had 100 participants (obtained through a university research pool and snowball sampling) submit a sample of their Facebook profile consisting of recent posts, recent, profile pictures, and the info page (now called About), and complete a personality assessment. Following this, all profile samples were removed of identifying information and coded by up to four independent coders for 45 of the 53 cues. Examples of cues included things self-generated information (use of positive affect in a status update), other-generated information (number of likes or comments on a post) and system-generated information (total number of Facebook friends). The remaining 8 cues were analyzed using Diction (Hart & Carroll, 2011) for word and character counts. Finally, to complete the lens model process, up to 35 observers viewed the 100 profiles to evaluate them for the same personality traits self-reported. Here is a brief summary of the diagnostic cues found in the main study:

- Extraversion: higher number of Facebook friends, use of positive affect in status updates, friends in pictures
- Openness: listing books, music on info page, sharing media in status updates, posting political status updates
- Conscientiousness: friends' agreement with posts
- Agreeableness: post fewer updates, less media and news posted
- Neuroticism: No cues

Strangers accurately estimated extraversion, agreeableness, and conscientiousness.

A separate study (Pennington & Hall, 2014) was conducted to examine whether humor orientation could be accurately estimated on Facebook. Humor orientation was

estimated at rates similar to extraversion, and there were several places on the profile page that reflected users' sense of humor: a lack of political talk and increased relational talk in status updates, likes on status updates, and use of jokes in pictures, quotes, and status updates.

A final study explored the idea that there is a tradeoff between honesty and self-promotion on Facebook that can be mapped onto user's conscientiousness and self-monitoring (Hall & Pennington, 2013). This study explored the idea that Facebook users were actively using their profile to construct a self-image that would favorably be evaluated with friends, and Facebook could be used to promote certain aspects of a person's personality.

We remain interested in several outstanding questions, such as, have changes in Facebook impacted the accuracy of personality judgments, and/or the degree to which users are actively constructing their profile page? New algorithms enacted by Facebook could be making it harder to detect personality traits. However, since much of the reviewed research studied earlier versions of Facebook, it's conceivable that people's ability to accurately judge others will go down as a consequence of these changes. Since 2011, Facebook has changed how and when users see other people's activity. At the time our data was collected, users saw every action – from likes to changes in personal history – their friends took. Now, those actions can be viewed in a small box in the upper right-hand corner of the page, making the actions less apparent. Today, the posts on Facebook's most prominent feature, the newsfeed, are based on an algorithm that takes into account how recent the post is, how many people like it or have commented on it and if the user has frequently interacted with the person making the post, which affects how much and what type of information users see from one another. Another change is the kind of information that was shown on the About page. At one point, Facebook users were able to list their favorite bands, books and movies. Those who did so tended to have open personalities. Now, Facebook asks users to choose from a list of options, which is a passive step versus an active one. An open person was able to construct their personality through the process of making choices. Facebook has taken away agency and replaced it with algorithms.

References:

- Hart, R. P., & Carroll, C. (2011). *DICTION 6.0* [Computer software]. Retrieved from <http://www.dictionsoftware.com/index.php>.
- Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. *Journal of Computer-Mediated Communication, 13*, 531-549. doi:10.1111/j.1083-6101.2008.00409.x
- Walther, J. B., & Parks, M. R. (2002). Cues filtered out, cues filtered in: Computer-mediated communication and relationships. In M. L. Knapp & J. A. Daly (Eds.), *Handbook of interpersonal communication* (pp. 529-563). Thousand Oaks, CA: Sage.

For further reading please see the following:

- Hall, J. A., & Pennington, N. (2013). Self-monitoring, honesty, and cue use on Facebook: The extraversion and conscientiousness tradeoff. *Computers in Human Behavior, 29*, 1556-1564. doi:10.1016/j.chb.2013.01.001.
- Hall, J. A., & Pennington, N. (2012). What you can really know about someone from their Facebook profile (and where you should look to find out). In C. Cunningham (Ed.), *Social networking and identity management*. Lanham, MD: Lexington Books.
- Hall, J. A., Pennington, N., & Lueders, A. (2014). Impression creation and formation on Facebook: The lens model approach. *New Media & Society, 16*, 958-982. doi:10.1177/1461444813495166.

Pennington, N., & Hall, J. A. (2014). An analysis of humor orientation on Facebook: A lens model approach. *Humor: International Journal of Humor Research*, 27, 1-21. doi:10.1515/humor-2013-0053.

Tool to Produce Desired Cue-Cue Intercorrelations in a Stimulus Set

Robert M. Hamm

*Clinical Sciences Division, Department of Family and Preventive Medicine,
University of Oklahoma Health Science Center, Oklahoma City, Oklahoma,
USA*

Contact: Robert-hamm@ouhsc.edu

A perennial issue in research using the Brunswick Lens model approach is control of the cue-cue intercorrelations in the stimuli (e.g., Brehmer, 1974) so they may be representative of the structure of the environment. Decades ago there were some programs distributed which offered to help researchers produce stimuli with correlated cues, but they have long disappeared from my computer. So I have written an Excel spreadsheet that gives some control over this (CueCorrels.xlsx at www.fammed.ouhsc.edu/robhamm/cdmcalc.htm).

The Brunswick approach may be contrasted with the approach of experimental psychology or policy capturing research. Experimental psychology emphasizes control of stimuli, and hence it is natural to arrange stimulus cues to vary independently so we may most efficiently identify how much impact each cue has on people's judgments. In contrast the Brunswick approach emphasizes that the actual ecology that the judge would have learned in, and takes judgment-guided action in, should be reflected in the stimulus set of a study. The distinction was shown by Phelps and Shanteau (1978), who found that swine judges can discriminate using 11 different cues if the stimuli are constructed combinations of cue values, varying independently. However, when the swine were presented as photographs, with naturally occurring cue intercorrelations, the study could only demonstrate the experts' use of about 3 cues.

The argument that the cue-cue intercorrelations in a study's stimuli should be representative of the ecology's cue intercorrelations is that the design with uncorrelated cues does not let us observe when the ecology can be managed with a variety of cue utilization strategies, because the cues are partially intersubstitutable (Dhimi, Hertwig et al., 2004). But what is the appropriate cue-cue intercorrelation? For the swine judges, ongoing breeding changes the intercorrelations of swine features over time. Similarly, for the task of judging how ill patients in the emergency room are (Yang, Thompson et al., 2013), the cue-cue intercorrelations in a big city hospital might differ from those in a small town, a refugee camp, or a combat medic tent.

There are two main approaches to constructing stimulus sets with desired cue-criterion and cue-cue intercorrelations: to sample the relevant ecology, or to construct the stimulus set to have desired intercorrelations (based on prior knowledge). If one has a representative set of cases from the target ecology, one may draw a sample to use as the stimulus set. In variants I am familiar with, such a sample was based mainly

on a convenient feature of the stimuli (short stretches of Colorado highways; Hammond, Hamm et al., 1987), or was drawn randomly controlling only the criterion (emergency department patients who had bad outcomes; Yang, Thompson et al., 2013). Pitfalls were encountered in each of these studies. The design of the study of engineers' judgments of highways required film strips of the highway segments as well as data about their safety and vehicle capacity. To make the task acceptable, we sought the shortest segments in the highway department data base; and the cue intercorrelations followed from that selection. We did not independently verify that the intercorrelations in the stimulus set were similar to the correlations in the set of all Colorado highways. With the study of emergency room patients, it was possible to compare the stimulus set cue intercorrelations with those of the larger patient data set, and they differed; our discussion of the differences is "in preparation".

The other approach is to construct stimulus sets with the desired cue intercorrelations in a top down manner, such as by stating the desired correlations, constructing multiple sets expected to have those intercorrelations, and selecting a set where the observed correlations are close enough to the desired targets. There are two strategies: to generate the stimuli in one process; or to generate the stimuli separately for each of two classes, such as "sick" or "well," with different cue correlations (see discussions of class-conditional dependence and of spectrum effects in Hamm & Beasley, 2014) and then combine the sets. The Excel spreadsheet uses the first approach.

To use the spreadsheet, specify the domain (how many cues), the number of stimuli needed, and the number of ordered categories for each cue and the criterion. With each change of any cell entry whatsoever, or press of the F9 key, a complete new set of stimuli is randomly generated and summarized for your inspection. The stimuli are displayed both as the continuous random variable and categorized into as many levels as you intend to use. Additionally, summary statistics (cue correlations and others) are generated. Copy and preserve the good stimulus sets, those whose actual correlations are close to what is desired.

The spreadsheet allows control of two types of correlation: cue-criterion (related to the relative weights in the ecology's regression model) and cue-cue. These desired correlations are written into selected cells in a "control" correlation matrix. If only cue-criterion correlations are needed, the set of stimuli with the desired cue-criterion correlations is created with these steps:

Create the criterion – simply a random variable: $Criterion_i = rand()-0.5$, which produces a uniform distribution with mean of 0.

Create each correlated cue, as a mix between the criterion variable and a new random draw:

$$Cue_{ki} = r(Cue_k, Criterion) * Criterion_i + \sqrt{1 - r(Cue_k, Criterion)^2} * (rand() - 0.5)$$

Elements of the formula are familiar from the lens model equation (Stewart 1976), and can be found at www.sitmo.com/article/generating-correlated-random-numbers/. We must recognize that when correlations are specified between cues and the criterion, we can expect the cues to be intercorrelated

$$r(Cue_j, Cue_k) = r(Cue_j, Criterion) * r(Criterion, Cue_k)$$

by basic path analysis principles. Though stimulus sets could be generated with cues that are not correlated, this is very unlikely. When we choose to specify a cue-cue correlation in addition to a cue-criterion correlation, that correlation is going to be ON TOP OF the expected cue-cue correlation that follows from the fact that each cue is correlated with the criterion. The spreadsheet generates the cue-cue correlations first, using a formula like the above but with the “top cue in the factor” substituted for the “criterion”. Then a second pass combines these intercorrelated cues with the criterion, for a stimulus set that has both cue-criterion and cue-cue correlations.

The spreadsheet offers separate sheets allowing you to have no extra cue-cue correlations, to have one factor of cue-cue correlations, or to have three factors. The work to allow different cue-cue correlations for different classes of the criterion, i.e., class conditional dependence (Hamm & Beasley, 2014) is not yet available in a debugged spreadsheet.

There are other ways to control the correlations among a set of cues. One method, offering complete control of every single pairwise correlation (rather than generating, evaluating, and selecting) uses matrix manipulation and Cholesky decomposition (<http://www.sitmo.com/article/generating-correlated-random-numbers/>). I think that the approach here makes the user more aware of the meanings of sampling and cue-cue correlations. Watching the variation in the cue intercorrelations as the data sets are randomly produced trains one’s understanding of these issues.

This work has been informed by discussions with Jonathan Nelson and Jana Jarecki at the Max Planck Institute’s Center for Adaptive Behavior and Cognition in Berlin, and with Pádraig Mac Neela, Sinead Conneely, and Chris Dwyer at the National University of Ireland in Galway.

References:

- Brehmer, B. (1974). The effect of cue intercorrelation on interpersonal learning of probabilistic inference tasks. *Organizational Behavior & Human Performance*, 12(3), 397-412.
- Dhimi, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, 130(6), 959-988.
- Hamm, R. M., & Beasley, W. H. (2014). The balance beam metaphor: A perspective on clinical diagnosis. *Medical Decision Making*, 34(7), 841-853.
- Hammond, K. R., Hamm, R. M., Grassia, J. L., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions on Systems, Man, and Cybernetics*, 17, 753-770.
- Phelps, R. H., & Shanteau, J. (1978). Livestock judges: How much information can an expert use? *Organizational Behavior and Human Performance*, 21, 209- 219.
- Stewart, T. R. (1976). Components of correlation and extension of the lens model equation. *Psychometrika*, 41(1), 101-120.
- Yang, H., C. Thompson, C., Hamm, R. M., Bland, M., & Foster, A. (2013). The effect of improving task representativeness on capturing nurses' risk assessment judgments: A comparison of written case simulations and physical simulations. *BMC Medical Informatics and Decision Making*, 13, 62.

Information Aggregation in Online Communities

Steffen Herm

*Hochschule für Technik und Wirtschaft (University of Applied Sciences),
Berlin,
GER*

Contact: herm@htw-berlin.de

Our article entitled “When the crowd evaluates soccer players’ market values: Accuracy and evaluation attributes of an online community” has recently been accepted for publication by the *Sport Management Review* – one of the finest journals in the field of sport management. This paper is the result of a fruitful cooperation with Dr. Hans-Markus Callsen-Bracker, my friend and former colleague at Technische Universität Berlin, and Henning Kreis, my friend and professor of marketing at the Freie Universität Berlin. Since our conceptual model builds on the work of Egon Brunswik and Berndt Brehmer we received the honor of an invitation to contribute to this newsletter of the Brunswik Society.

Sir Francis Galton discussed the idea of “collective wisdom” when he asked “the crowd” to gauge the weight of an ox. He successfully employed a democratic principle (an equal say) to arrive at a precise estimation. Since Galton, the concept of “collective wisdom” has nourished a rich and interdisciplinary stream of research and popular science literature. Although the concept of “crowdsourcing” has been around for a long time, the advent of the Internet has opened up many new possibilities. Self-selected members of “the crowd” work together on problems such as improving algorithms, translating websites, and evaluating soccer players – quickly and worldwide.

The evaluation of a monetary value of athletes, e.g., in order to prepare negotiations about salaries or transfer fees (in the case of an athlete’s transfer from one club to another), has become a major challenge for managers of professional sports teams. In the context of German soccer, aficionados have built a large online community that evaluates professional soccer players’ market values. The community has become the main source for reporting market values in the media and has a strong impact on the (sports) economy.

We introduce an adaptation of Brunswik’s lens model for studying organizational processes in such an online community that solve complex tasks (such as the evaluation of human capital). More specifically, our research investigates how – with respect to the organizational principle – and how well – with respect to the accuracy of the evaluation output – an online community estimates professional soccer players’ market values, and which attributes are most important for the evaluations made by the crowd. We also contribute to the literature on the superstar phenomenon by reconsidering various well-known attributes as being external variables and not talent variables.

Our research shows that a community does not need democratic principles or structured approaches. Instead, we show that a principle that we call “the judge principle” – a more flexible system combining crowdsourcing of information with a final evaluation by expert decision makers – exhibits convincing results. For instance, we analyze evaluations of market values for professional soccer players resulting from a well-established German online community that uses the “judge principle” and find its predictions of actual transfer fees of impressive quality.

References:

- Brehmer, B., & Hagafors, R. (1986). Use of experts in complex decision making: A paradigm for the study of staff work. *Organizational Behavior and Human Decision Processes*, 38, 181-195.
- Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: The University of Chicago Press.
- Galton, F. (1907). Vox populi. *Nature*, 450.
- Herm, S., Callsen-Bracker, H.-M., & Kreis, H. (2014). When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review*, <http://dx.doi.org/10.1016/j.smr.2013.12.006>.

**A Lens Model Approach to Accurate Personality Judgments:
The Case of Neuroticism and Situational Relevance**

Sarah Hirschmüller & Boris Egloff
*Johannes Gutenberg University Mainz,
GER*

Stefan C. Schmukle
*University of Leipzig,
GER*

Steffen Nestler & Mitja D. Back
*University of Münster,
GER*

Contact : hirschmu@uni-mainz.de

Even before we interact with people (zero acquaintance; Kenny & West, 2008), we form first impressions of their personality based on observable behavior. These personality inferences are consequential and may influence our thoughts and behaviors toward others. In research on the accuracy of first impressions, the lens model framework (Borkenau & Liebler, 1992; Brunswik, 1956; also see Nestler & Back, 2013) has been extensively applied to study fundamental questions of personality judgment: How accurate are first impressions and what processes underlie more or less accurate personality judgments? Typical situations based on which targets' personality traits have been judged by strangers include brief interactions, videotaped situations, or photographs in laboratory settings. For the Big Five personality traits, meta-analytic evidence across a large number of zero-acquaintance studies shows that accuracy is highest for extraversion whereas neuroticism appears to be extremely difficult to judge (Connelly & Ones, 2010; Connolly, Kavanagh, & Viswesvaran, 2007).

In the case of neuroticism, this pattern of empirical results is surprising. Neuroticism is related to a large array of negative interpersonal outcomes such as dissatisfaction, conflict, and relationship dissolution (cf., Ozer & Benet-Martínez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). It should be advantageous from an evolutionary perspective (Haselton & Funder, 2006) to be able to detect an individual's level of emotional instability, even when (still) unacquainted. How then, can the low convergence between targets' actual trait value and strangers' judgments of the trait be explained? According to Brunswik's (1956) lens model, a target's personality trait can be inferred accurately only when (a) the trait is expressed via trait-relevant and observable cues (cue validity; e.g., a neurotic target person behaves in an observable nervous manner) and (b) the perceiver observes and correctly utilizes these cues for the personality judgment (cue utilization; e.g., the perceiver notices and judges the target in accordance with the nervous behaviors). With regard to a realistic interactionist view of personality (Kenrick & Funder, 1991), valid cues can be expressed (and thus potentially used by perceivers) in trait-relevant situations only when interindividual differences concerning the specific trait are actualized and translated into manifest behaviors. In prior research, neuroticism has been shown to

be related to a heightened sensitivity to social threats, a hyperreactivity to uncertain feedback as well as interpersonal stressors (Hirsh & Inzlicht, 2008; Denissen & Penke, 2008; Zautra, Affleck, Tennen, Reich, & Davis, 2005), and consequently to poorer self-control in public social contexts (Uziel & Baumeister, 2012). An evaluative public social situation crucial for the formation of new social bonds compared to situations less public as well as situations by oneself can thus be assumed to be especially relevant with regard to the trait of neuroticism.

In our own research (see Hirschmüller, Egloff, Schmukle, Nestler, & Back, 2014), we aimed to reexamine the low convergence between people's actual trait value of neuroticism and neuroticism judgments of strangers reported in prior research by implementing a highly trait-relevant situation as well as three less relevant situations. Further, we wanted to analyze the behavioral processes underlying the level of zero-acquaintance accuracy for these situations using lens model analyses. To this end, fifty participants were videotaped in an evaluative public social (i.e., socially stressful) situation during brief self-introductions to fellow students at the time of first meeting. The same participants were videotaped in three further less trait-relevant situations during a subsequent individual laboratory session (a welcoming situation by one experimenter, a dyadic waiting situation, and an individual waiting situation). Based on these short video sequences, four independent groups of unacquainted observers judged participants' neuroticism. We used an aggregate of neuroticism self-reports and informant reports as the accuracy criterion. Moreover, we applied a single-perceiver approach throughout all our analyses as the commonly employed aggregate-perceiver approach tends to overestimate results due to averaging across individuals' judgments prior to analyses (cf. Hall & Bernieri, 2001; Hall, Bernieri, & Carney, 2005; Hirschmüller, Egloff, Nestler, & Back, 2013; Nestler, Egloff, Kүfner, & Back, 2012).

Our results showed that neuroticism judgments were significantly more accurate for the most trait-relevant introductory situation compared with the other three situations. Using lens model analyses, this finding could be explained: Only in the socially stressful situation did neuroticism predict both visual nervousness and vocal nervousness (cue validity), both of which predicted neuroticism judgments by unacquainted perceivers (cue utilization). Thus, our findings suggest that relevant situations can indeed elicit observable behaviors that are valid cues to neuroticism which can be used by strangers to accurately judge interindividual differences in neuroticism. We wish to encourage researchers to apply and extend the lens model approach to further elucidate more or less accurate interpersonal judgments by using a single-perceiver approach and realizing "studies in which individuals are each placed into or observed in each of a range of different situations" (Funder, 2009, pp. 124-125).

References:

- Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero acquaintance. *Journal of Personality and Social Psychology*, 62, 645-657.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136, 1092-1122.
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, 15, 110-117.
- Denissen, J. J. A., & Penke, L. (2008). Neuroticism predicts reactions to cues of social inclusion. *European Journal of Personality*, 22, 497-517.

- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, *102*, 652-670.
- Hall, J. A., & Bernieri, F. (Eds.). (2001). *Interpersonal sensitivity: Theory and measurement*. Mahwah, NJ: Erlbaum.
- Hall, J. A., Bernieri, F., & Carney, D. R. (2005). Nonverbal behavior and interpersonal sensitivity. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook in nonverbal behavior research* (pp. 237–281). Oxford: Oxford University Press.
- Haselton, M. G., & Funder, D. C. (2006). The evolution of accuracy and bias in social judgment. In M. Schaller, D. T. Kenrick, & J. A. Simpson (Eds.), *Evolution and social psychology* (pp. 15-37). New York: Psychology Press.
- Hirschmüller, S., Egloff, B., Nestler, S., & Back, M. D. (2013). The dual lens model: A comprehensive framework for understanding self-other agreement of personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, *104*, 335-353.
- Hirschmüller, S., Egloff, B., Schmukle, S. C., Nestler, S., & Back, M. D. (2014). Accurate judgments of neuroticism at zero acquaintance: A question of relevance. *Journal of Personality*.
- Hirsh, J. B., & Inzlicht, M. (2008). The devil you know: Neuroticism predicts neural response to uncertainty. *Psychological Science*, *19*, 962-967.
- Kenny, D. A., & West, T. V. (2008). Zero acquaintance: Definitions, statistical model, findings, and process. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp. 129-146). New York: Guilford Press.
- Kenrick, D. T., & Funder, D. C. (1991). The person-situation debate: Do personality traits really exist? In V. J. Derlega, B. A. Winstead, & W. H. Jones (Eds.), *Personality: Contemporary theory and research* (pp. 149-174). Chicago: Nelson-Hall.
- Nestler, S., & Back, M. D. (2013). Applications and extensions of the lens model to understand interpersonal judgments at zero acquaintance. *Current Directions in Psychological Science*, *22*, 374-379.
- Nestler, S., Egloff, B., Kűfner, A. C. P., & Back, M. D. (2012). An integrative lens model approach to bias and accuracy in human inferences: Hindsight effects and knowledge updating in personality judgments. *Journal of Personality and Social Psychology*, *103*, 689-717.
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401-421.
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*, 313-345.
- Uziel, L., & Baumeister, R. F. (2012). The effect of public social context on self-control: Depletion for neuroticism and restoration for impression management. *Personality and Social Psychology Bulletin*, *38*, 384-396.
- Zautra, A. J., Affleck, G. G., Tennen, H., Reich, J. W., & Davis, M. C. (2005). Dynamic approaches to emotions and stress in everyday life: Bolger and Zuckerman reloaded with positive as well as negative affects. *Journal of Personality*, *73*, 1511-1538.

**Judgment Studies and Economics
of Information Exchanges in Healthcare**

Christine Huttin

ENDEPUSresearch, Inc., USA,
University Aix Marseille, FRA

Contact: chris.huttin@comcast.net

The move from a market economy to an exchange economy creates new needs for information systems to complement conventional pricing methodologies used to analyze market adjustment of supply and demand.

Cost and price measures in current methodologies for valuation of goods and services do not seem sufficient to understand economic choices and the way more information can influence decision making processes. The Brunswik society can largely contribute to this research area, since judgment studies especially in medical sciences have advanced exploration of decision making processes and understanding of incoherencies in information processed for instance by physicians, other providers of care and patients.

With the recent increase of information exchanges on health insurance markets, for instance, a new trend of research is emerging in order to provide new cost concepts and associated measures can help to better understand the choices of patients and companies facing different options for insurance plans.

Recent research on cost informativeness analysis (e.g., Hamm & Scheid, 2014) for tests in oncology shows for instance how degrees of uncertainty for diagnostic choices are influenced by different levels of access to information on tests and discuss the results of a cost informativeness analysis differ from a cost effectiveness analysis.

Other research on cost awareness and cost consciousness of medical services with judgment studies (e.g., Huttin, 2014) identify the type of economic information in the verbal language of doctors with their patients which can be used and transformed into reliable metrics to better understand decision treatment or diagnostic shifts in critical decision points when economics interfere with clinical decision making; it can complement conventional price and copayment information.

Further collaboration among the researchers interested in new cost and price concepts and measures could probably help to build statistically reliable tests; the society could also benefit from current advances coming from accounting and financial research (Chen, Huang, & Zhang, 2014; Christensen & Qin, 2014) especially on the effect of more public information on market asymmetry, and especially principal-agent contracts, largely used to analyze health care insurance markets.

References:

Hamm, R. M., & Scheid, D. C. (2014). *Cost informativeness ratio as a guide for whether to purchase diagnostic information*, Communication at SMDM, Antwerpen, June 10, 2014

- Huttin, C. (2012). A rum model on physicians' choice sets and health care financing systems, Inter-country case study comparisons and data elements for model development. SMDM, Atwerpen, June 10, 2014 and Society for Judgment, US, 2012.
- Christensen, P. O., & Qin, Z. (2013). Information and heterogeneous beliefs: Cost of capital, trading volumes and investor welfare, *Accounting Review*, 89(1), 209-242.
- Chen, Q., Huang, Z., & Zhang, Y. (2014). The effects of public information with asymmetrically informed short-horizon investors. *Journal of Accounting Research*, 52(3), 635-669.

Testing an Evolutionary Theory of Human Rhythm and Groove: Tapping Musicians' Implicit Knowledge through Egon Brunswik's Lens Model

Guy Madison

*Umeå University, Department of Psychology,
Sweden*

Contact: guy.madison@umu.se

Music often induces movement. Hearing the old masters in the concert hall or the latest dance music on the radio may both make us tap our feet or bob our head. This behavior tendency is called Groove, defined as the experience of wanting to move when hearing music. Starting out of mere curiosity about this phenomenon, from a musician's point of view, I have come to think that it is a human universal that might reflect an adaptive value of being able to synchronize actions with conspecifics. My initial goal, however, was simply to crack the code of how to play so as to induce as much groove as possible. To this humble end one needs to know, first, whether groove is personal or general. If a piece of music that makes me experience groove leaves another person indifferent there is little hope of finding this Holy Grail.

The common impression that some artists do a better job than others at inducing groove is correct, though: About 25 percent of the variability in listeners' ratings of groove could be explained merely by the differences amongst 64 examples of popular music (Madison, 2006). This music was sampled from commercially available audiograms, and was therefore unlikely to vary systematically in groove induction ability. Given the suboptimal stimulus sampling, individual differences in scaling behavior, and measurement error, 25 percent reflects a high degree of consistency across individuals. The next step was to explore which physical properties are associated with groove.

Starting with a naturalistic sample of real music, we measured a number of higher-order rhythmic properties in the sound signals of 100 commercially available music recordings, and related them to listeners' ratings of groove for each recording (Madison et al., 2011). We constructed a few dozen so-called descriptors of the audio waveform, essentially computational algorithms that focus on various aspects of the signal, from lower-order spectral properties to higher-order ones. For example, systematic micro-timing involves first analyzing the beat and metrical properties, then applying canonical time according to the meter, and finally measuring deviations between these and the actual signal events. Each descriptor yields a single parameter that expresses the magnitude of the property in question. The descriptors called beat salience, event density, fast metrical levels, and systematic micro-timing were

significantly correlated with the listeners' ratings. Specifically, beat salience reflects the number and loudness of sounds that occur on the beat, event density is the number and loudness of sound events per unit time, fast metrical levels is the number of metrical subdivisions that sound events are articulated in, and systematic micro-timing is the amount of temporal deviations from the metronomic (i.e., canonical) positions of the metrical grid that are recurrent and hence not random. All these descriptors were positively correlated with groove except micro-timing, which was negatively correlated (Madison et al., 2011). Having thus obtained ideas about the physical properties involved from the real music itself and from the listeners themselves, we would seem to be in a position to experimentally test their validity. But a multifaceted and highly complex phenomenon like music requires an even more careful approach to exhaust the possible physical correlates of groove. I tend to refer to the "black box" problem, meaning that music has so many simultaneous properties that it is practically impossible to control them unless it is simplified *reductio ad absurdum*. Consider at least two apparent problems.

First, a sample of real music might contain only a few or even no pieces actually intended to induce movement. There are certainly many other desirable aspects of music. In fact, derivative styles of what was originally dance music, such as contemporary jazz, samba, tango, and many other from the Latin and Black Atlantic diaspora, are today performed for listening exclusively.

Second, Groove may be trivially associated with a range of properties that are non-essential for the intention to produce groove. This is a variety of the confounding variables problem. For example, speech comprehension hinges on a range of variables under direct control of the speaker and listener, such as stresses, timing, and anticipation, but also on the quality of the voice, impaired hearing, and background noise, that would trivially affect the perception of any auditory signal. Similarly, a musical tradition that, for example, encompasses a focus on the lyrics, a smaller budget, and a desire for natural sounds is likely to simultaneously feature slow tempo, few instruments, and a narrow frequency spectrum. Now, inasmuch as the focus on lyrics might also render groove unimportant we cannot tell if this intention has any effect in and of itself, since it would be confounded with the other features that also lend the music less movement inducing: A slow tempo is difficult to accommodate to body movement (MacDougall & Moore, 2005), few instruments provide less opportunity for explicating fast metrical levels, syncopation, and rhythmic elaboration (Madison et al., 2011; Madison & Sioros, 2014; Sioros et al., 2014), and less low frequency range power fails to engage the vestibular system (e.g., Todd, 2001). In conclusion, "black box" phenomena require special care to avoid, on the one hand, reducing them to something out of their true nature, and, on the other hand, a number of interpretation and design problems related to confounding variables.

One approach to avoid these problems is to examine what musicians would do if asked to increase or decrease groove across a range of different musical structures. Brunswick's (1952; 1956) lens model is a potent tool for such situations, in particular because it takes the whole communication process into account so that we can confirm that their behavior has the intended effect on listeners' perception. In other words, we exploit musicians' experience and expressive skills directly by asking them to play a number of compositions with as much and as little groove as possible, to amend our exploratory examinations of pre-existing musical pieces with high ecological validity. The compositions were monophonic melodies, and the musicians

had to play all notes in the specified order in a steady, specified tempo. They were however allowed to add notes and to change the note values. These depleted conditions were intended to refine musicians’ strategies by focusing on the devices at hand, mainly timing, dynamics, and rhythmization. Four professional musicians performed 12 monophonic melodies; six simple melodies akin to children’s songs that were composed for this study, and six complex ones were adapted from jazz and rock style recordings. The musicians could hear and rehearse the melodies at home, and at the production session each musician individually first recorded all 12 melodies in a deadpan version similar to the one rehearsed. It was only after this that they were instructed to play the same melodies again with the intention of maximizing and minimizing groove. They played a professional Yamaha keyboard through which information about each keystroke was recorded as MIDI data. The 24 performances were subjected to both listener ratings and performance analysis, totaling 96 performances from the four musicians. Ten different performance parameters were computed from the MIDI data across each performance, namely event density, the magnitude of onset, offset, and duration micro-timing, and the proportions of 8th and 16th note onsets, offsets, and durations. Details about methods, analyses, and results are given in Madison and Sioros (2014).

Thirty non-musicians then rated each of the performances on how movement inducing it was on a scale from 0 (“not at all”) to 10 (“entirely”). Movement inducing was defined as “the sensation of wanting to move some part of your body in relation to some aspect of the music”. The ratings were entered through a slider on the computer screen, and ANOVA tests demonstrated significant effects of intention on Groove ratings.

Table 1 summarizes the lens model results from the dichotomous instruction to induce as much or as little groove as possible, over the 10 performance parameters, and to the 11-point groove rating scale.

Table 1
Lens model factors for simple, complex, and both types of melodies

Melody type	r_a	G	R_s	R_r
Simple	0.620	0.83	0.749	0.687
Complex	0.394	0.34	0.729	0.523
Both	0.489	0.78	0.621	0.490

Equation 1 shows that the communication achievement (r_a) is the product of the matching factor G, the multiple correlation between musicians’ (the senders) intention and performance parameters R_s , and the multiple correlation between performance parameters and the listeners’ (the receivers) ratings R_r plus an un-modelled component that consists of the correlation between the residuals in the regression models (C) and the residual variation of the models (Hursch, Hammond, & Hursch, 1964):

$$r_a = G R_s R_r + C \sqrt{(1 - R_s^2)} \sqrt{(1 - R_r^2)} \tag{1}$$

where G is the correlation between the predicted values for the sender and receiver regression models, and C is the correlation between the residual values for the sender and receiver regression models.

The results demonstrate successful communication for the simple melodies but relatively poor for the complex ones. A high G for the simple melodies indicates that performers and listeners share a common code. The two R factors indicate that the cues are not used fully consistently, in particular not by the listeners. Matching primarily limits the communication for the complex melodies, because the performers' cue utilization is almost equal to that for the simple melodies. The un-modelled component C varied from 0.37-0.44, which is rather low and indicates that we have not excluded cues that performers and listeners would have used consistently. Although not shown here, significant positive correlations between intention and performance cues were found for event density and proportions of 8th and 16th note onsets and 8th offsets (.38 - .57), whereas cue correlations for micro-timing were small and negative.

Overall, these results support those based on real music (Madison et al., 2011) by exhibiting a quite similar pattern of correlations between groove and physical properties. However, musicians' addition of 8th and 16th note onsets when none were specified in the score show that they intentionally syncopate to increase groove, that is, tend to play stressed notes on relatively weaker positions in the beat or metrical structure (Randel, 1986). Likewise, syncopes and other notes on weak metrical positions were sometimes moved to the beat when musicians decreased groove in the complex melodies. This is a very important qualification of the real music results, where the correlations with density and fast metrical levels could not discriminate between syncopation and the presence of other sounds on weak positions, such as perfectly metrical rhythmic patterns like those typically played by the shaker, hi-hat, tambourine, or rhythm guitar. The poor communication for the complex melodies was attributed to a ceiling effect resulting from their already busy structure, which leaved little room for adding syncopation.

Having thus confirmed, with two quite different approaches, that groove is positively associated with event density and syncopation, and negatively associated with micro-timing (although far from significantly so in the musician study), we proceeded to test these associations experimentally in subsequent studies (Davies, Madison, Silva, and Gouyon, 2013; Sioros, Miron, Davies, Gouyon, & Madison, 2014).

The adaptive perspective on the perception and production of rhythmical patterns presupposes that the abilities and behaviors involved are functional, and we have proposed that their function is to facilitate temporal prediction and synchronization (Madison & Merker, 2005; McNeil, 1995; Merker, Madison, & Eckerdal, 2009). One specific evolutionary scenario posits that producing loud signals by joint vocal exclamations make these signals reach farther and attract larger numbers of conspecifics (Merker, 1999). Several observations support a phylogenetic history. Music is a human universal (Pinker, 2002), and coordinated dance to rhythmically predictable music seems to occur in all cultures (Nettl, 2000). Experiencing rhythmic music is associated with pleasure (Madison, 2006; Todd, 2001; Witek et al., 2014) through activation of brain areas associated with reward and arousal (e.g., Blood & Zatorre, 2001). Finally, passive listening to music and rhythmic sequence activates motor system areas even for tasks without any reference to movement (e.g., Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007). Groove

would according to such a scenario constitute the motivational tendency for synchronization, and should conceivably be related to the signal's effectiveness for synchronization. Consistent with this, faster metrical levels were associated with greater synchronization accuracy (Madison, 2014).

In conclusion, the application of a Brunswikian lens model presented herein provided critical knowledge for extending previous findings into experimentally testable hypotheses. A series of cumulative designs have so far failed to falsify the functional theory of rhythm, which posits that synchronization is associated with an adaptive value (Merker et al., 2009; Madison et al., 2011). The findings are generally consistent with the idea that Groove reflects the behavioral tendency to engage in synchronization, as well as the utility of an auditory signal to facilitate precise synchronization.

Acknowledgements

This research was supported by grant P2008:0887 to Guy Madison from the Bank of Sweden Tercentenary Foundation.

References:

- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, USA*, 98, 11818-11823.
- Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.
- Brunswik, E. (1956). *Perception and the representative design of experiments*. Berkeley: University of California Press.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18, 2844-2854.
- Davies, M., Madison, G., Silva, P., & Gouyon, F. (2013). The effect of micro-timing deviations on the perception of groove in short rhythms. *Music Perception*, 30, 498-511.
- Grahn, J. A., & Brett, M. (2007). Rhythm perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19, 893-906.
- Hirsch, C. J., Hammond, K. R., & Hirsch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Review*, 71, 42-60.
- MacDougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology*, 99, 1164-1173.
- Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, 24, 201-208.
- Madison, G. (2014). Sensori-motor synchronisation variability decreases with the number of metrical levels in the stimulus signal. *Acta Psychologica*, 147, 10-16.
- Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with some low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1578-1594.
- Madison, G., & Merker, B. (2005). Timing of action during and after synchronization with linearly changing intervals. *Music Perception*, 22, 441-459.
- Madison, G., & Sioros, G. (2014). What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Frontiers in Psychology*, 5, 894. doi: 10.3389/fpsyg.2014.00894
- McNeil, W. H. (1995). *Keeping together in time. Dance and drill in human history*. Cambridge, MA: Harvard University Press.
- Merker, B. (1999). Synchronous choring and the origins of music. *Musicae Scientiae, Special issue, 1999-2000*, 59-73.
- Merker, B., Madison, G., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45, 4-17.
- Nettl, B. (2000). An Ethnomusicologist contemplates universals in musical sound and musical culture. In N. L. Wallin, B. Merker, & S. W. Brown (Eds.), *The origins of music* (pp. 463-472). Cambridge, MA: MIT Press.
- Pinker, S. (2002). *The Blank Slate. The Modern Denial of Human Nature*. London: Penguin.

- Randel, D. M. (1986). *The Harvard Dictionary of Music*. (4^{ed.}) Cambridge, MA: Belknap Press of Harvard University Press.
- Sioros, G., Miron, M., Davies, M., Gouyon, F., & Madison, G. (2014). Syncopation creates the sensation of groove in synthesized music examples. *Frontiers in Psychology*, 5, 1036. doi: 10.3389/fpsyg.2014.01036.
- Todd, N. P. M. (2001). Evidence for a behavioral significance of saccular acoustic sensitivity in humans. *The Journal of the Acoustical Society of America*, 110, 380-390
- Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, 9, e94446.

News from the Past Year

Jeryl L. Mumpower
U. S. National Science Foundation,
on leave from Texas A&M University,
USA

Contact: jmumpowe@nsf.com (jmumpower@tamu.edu)

For the past two years, I have been on leave from my position at Texas A&M University and have served as Division Director for Social and Economic Sciences at the U.S. National Science Foundation. I am now beginning my third and final year in this role. This continues to be an interesting experience and I believe I am engaged in important, valuable work. It has certainly affected my ability to get research done, but I continue to try to be as active as my responsibilities permit.

Gary McClelland (now emeritus) at the University of Colorado, Boulder, and I published a paper in *Judgment and Decision Making* entitled "A Signal Detection Analysis of Racial and Ethnic Disproportionality in the Referral and Substantiation Components of the U.S. Child Welfare Services System." We also presented some follow-on work entitled "Modeling Judgment and Decision Processes at the System Level," which we presented at the Decision Making Bristol 2014 Conference, UK.

My colleagues at Texas A&M University, Arnie Vedlitz, Xinsheng Liu, and I just completed a paper, now in review, entitled "Psychometric and Demographic Predictors of the Perceived Risk of Climate Change and Preferred Resource Levels for Climate Change Management Programs." Finally, Tom Stewart, Jim Holzworth and I continue work on our next paper stemming from our research program investigating how people make selection and detection decisions (e.g., how they to decide whether to hire someone or whether a patient has a disease) in the face of uncertainty and different feedback conditions.

References:

- Mumpower, J. L., & McClelland, G. H. (2014). A signal detection theory analysis of racial and ethnic disproportionality in the referral and substantiation processes of the U.S. child welfare services system. *Judgment and Decision Making*, 9(2), 114-128.

Probabilistic Functionalism and Systematic Error in the Wisdom of Crowds

Ulrik W. Nash

*Strategic Organization Design, SDU,
Denmark*

Contact: uwn@sam.sdu.dk

I am grateful to the Brunswik Society Newsletter for giving me this opportunity to explain the content of my recent paper in PLOS ONE (Nash, 2014), which was heavily inspired by the ideas of Egon Brunswik.

My study of Brunswik has made clear for me something you all know. Brunswik was an exceptional scientist. Indeed, it is no stretch to say he was so pioneering, that some of the ideas he communicated still remain unexplored today, despite their obvious importance. One such idea is the relation between the information environment, and the distribution of judgments formed by many people (1956). Few scientists had even contemplated the idea of judgment distributions before Brunswik, but one notable exception was Galton (1907), which perhaps is unsurprising, given the profound level of inquisitiveness both of these thinkers had.

Perhaps you already know the story of Galton's visit to Plymouth 1906, but if you do not, please allow me to tell it here. The story is important for understanding the investigation I conducted.

Upon visiting Plymouth, Galton attended the West of England Fat Stock and Poultry Exhibition, and while there, noticed a contest where the objective for participants was to guess the weight of an ox, after it had been slaughtered and dressed according to standard practice. With the aim of investigating the trustworthiness of popular judgment, Galton asked for, and was given access to, all of the 787 legitimate tickets that were submitted, which he subsequently examined. His results appeared in Nature the following year. In Galton's own words (1907, p. 450), his examination was "small but to the point". He simply arranged the submitted estimates in order of magnitude and identified the median, which he viewed as vox populi, that is to say, the single voice of contestants.

Galton made two important observations. The first observation concerned the median of estimates submitted by contestants. It was correct within a very small margin of the actual meat equivalent weight, and today this observation is recognized as an instance of a general phenomenon, which has become known to the general public as the Wisdom of Crowds, following Surowiecki's (2004) bestseller, which describes it.

The second observation Galton published in Nature on that occasion has largely gone unnoticed, but relates very much to what Brunswik would later examine. What Galton (1907, p. 450) also discovered was that estimates provided by contestants were not distributed symmetrically around the median, nor were the tails anything like that of Normal distributions. They were much wider. Baffled by what he

had seen, Galton left an intriguing conjecture for other scientists to examine (1907, p. 451):

I do not have sufficient knowledge of the mental methods followed by those who judge weights to offer a useful opinion as to the cause of this curious anomaly. It is partly a psychological question ... Also the anomaly may be partly due to the use of a small variety of different methods, or formulae.

My study of probabilistic functionalism helped me address Galton's hypothesis. Indeed, Brunswik's idea of probabilistic functionalism appears central to answering this puzzle, although so too were Thurstone's (1927) ideas about psychophysics.

In my work at Strategic Organization Design, simulation is my primary methodology, and my early attempt to program, judgment competitions of the kind Galton had observed provided excellent results. I simply endowed agents with regression models, and provided some with greater ability to utilize cues than others. Then, after providing agents numerous periods of learning about the correlation between oxen weight, and various perceptible oxen parts, I simply asked my agents to guess what the next exhibited ox weighed, and then proceeded to replicate Galton's study, not only for one competition, but for thousands. My findings were shared at Universitat Pompeu Fabra in 2010, where, incidentally, Professor Hogarth was in the audience.

My simulation had revealed numerous things. First, the frequency of skewed judgment distributions across competitions was high. Indeed, the frequency was much higher than what would be expected if judgment distributions were Normal. Second, when I correlated skewness with the error of the mean or median judgment, I observed clear negative correlation. In other words, not only had my code successfully captured Galton's second observation, but my simulation model clearly predicted an important relation to the Wisdom of Crowds, that is to say, Galton's more famous primary observation. It appeared judgment distributions were shaped by cues and the ability of my agents, becoming cues themselves for collective intelligence.

Unfortunately, I was finding it difficult to understand exactly how the ability of my agents, and their use of cues, was creating the effects I observed. The use of regression by my agents was hiding something crucial.

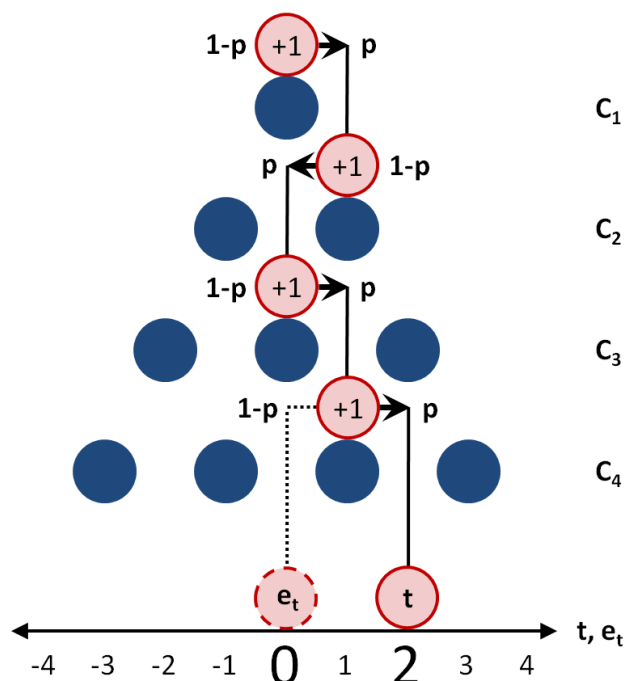
The AQ Model of Intuitive Judgment:

I soon made progress, however, after refreshing my knowledge of skewness, and seeing Galton's imminent probability device, the Quincunx, in a new light. I was naturally aware of its basic principle, and its purpose: the device was created to demonstrate the central limit theorem, and how the binomial distribution would tend towards the Normal distribution as an increasing number of balls passed the many layers of pins contained in the contraption. What I had not thought about before, however, was what happens when individual balls move past pins with uneven probability. When the probability of moving left, is different from the probability of moving right, the distribution of balls becomes skewed.

Then it dawned on me: „What if balls moving though the system of pins were conceived as individual judgments, and layers of pins as cues?“ When that simple idea

is pursued, the distal object can be computed by an 'attractor ball' moving around pins in the correct way, that is to say, in accordance with each of the multiple cues. For example, if the particular cue associates positively with the distal object, the attractor ball moves right at the corresponding layer of pins, while moving left when the association is negative.

The result was not exactly Galton's original Quincunx, but an extension. Galton's version has no attractor ball, and the probability of left movement, or right movement, is identical for all layers. In the Augmented Quincunx (AQ) model, however, the probability of following the attractor ball at every pin captures the individual judge's ability to categorize proximal cues correctly. Indeed, if this probability equals one, the judgment and the distal object become identical, but when it does not, something else happens entirely. Indeed, when individuals are fallible, yet categorize cues better than chance, judgment distribution skewness emerges, except only when the distal object is extreme compared to its typical value. That was the point regression had prevented me from seeing. Figure 1 below illustrates the AQ Model, while Figure 2 highlights the basic relation between it, and simple linear regression.



- t:** extremeness compared to what is typical.
- p:** probability of categorizing information correctly.
- e_t:** judgment about extremeness.

Figure 1. Although quite simple, the AQ model captures the probabilistic relation between our inferences about how unusual situations are, and what actually is, which is argued to originate from an uncertain cognitive process of categorizing information contained in cues, *C*. In the AQ model, rows of pegs represent cues, while balls falling through the system into one of various compartments represent the probabilistic categorization of these cues. Extremeness of the distal object, *t*, is computed by the distinct path taken by an attractor ball around pegs in the correct way.

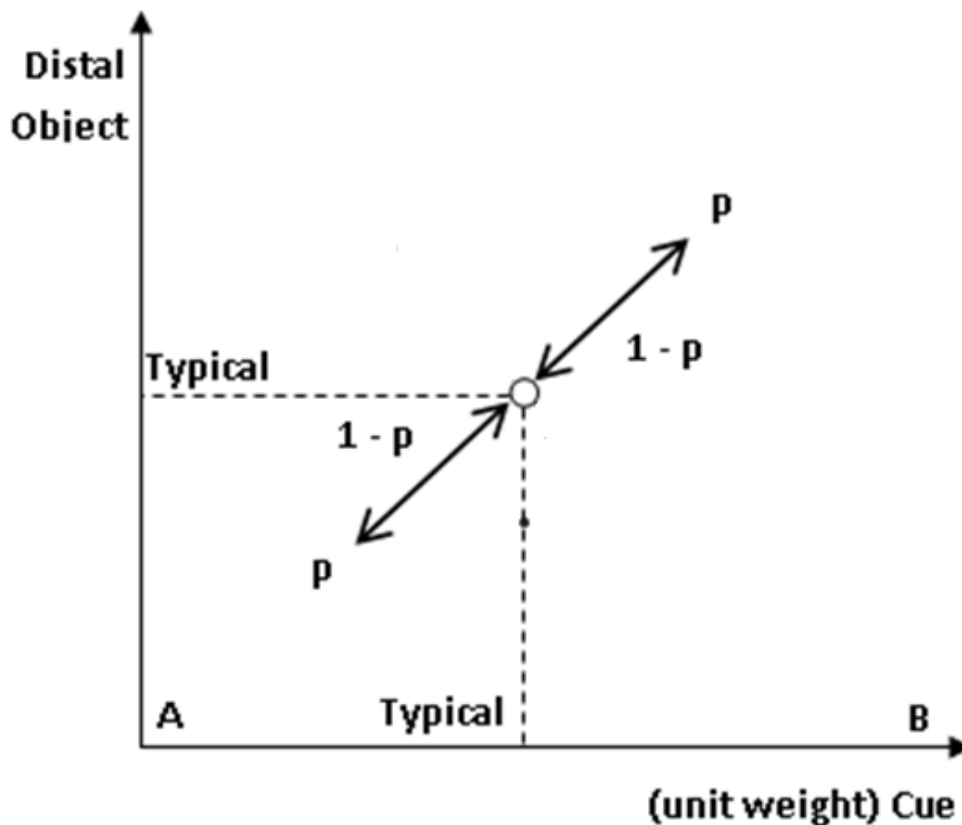


Figure 2. The AQ model assumes judges anchor on the intersection of the expected value of the distal object, and the expected value cues, and adjust away from that point by categorizing cues as being either greater or smaller than typical. The resemblance to standard regression is thereby quite clear, given that lines of regression pass through this point per definition. However, while the line of regression is placed on the history of data according to least squares, the AQ model captures the adjustment, judgment by judgment. Moreover, the adjustment is probabilistic to a degree determined by the judge's ability to categorize cues as having less (A) or more (B) information value than typical.

I will not describe further details of my study here, but hope this introduction to my examination has created sufficient interest for readers to visit PLOS ONE and learn in greater detail the role Brunswik has played in Galton's conjecture on the distribution of judgments finally being addressed. The supplementary materials to my PLOS ONE article includes an application, which is useful for visually discovering the AQ Model and its predictions.

References:

- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Univ of California Press.
- Galton, F. (1907) *Vox Populi*. Nature, 75 450-451 (March 7, 1907).
- Nash, U. W. (2014). *The curious anomaly of skewed judgment distributions, and systematic error in the Wisdom of Crowds*. PLOS ONE.
- Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*, Doubleday.
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34:273-286.

Recycling is Good!

Anna-Carin Nordvall

Umeå School of Business and Economics, Umeå University,
Sweden

Contact: anna-carin.nordvall@usbe.umu.se

My encounter with Brehmer has only been through his publications I eagerly read during my time as a doctoral student. In the dusty university library a hot summer day in 2000, I read “The psychology of linear judgment models” for the first time and I was hooked. Now, more than ten years later when I supervise my own doctoral students, his work still has a high impact on me.

The present study investigates whether there is any difference in learning speed, performance and knowledge representation (*cue abstraction* - knowledge representation of abstracted cue weights of the bug are added together to make a judgment or *exemplar memory* - individuals make judgments by retrieving similar exemplars from long-term memory) depending on whether the learning is made by written verbalization or reading verbalized rules. In the experiment three different groups were compared in performance (verbalized group, learning group and control group). The verbalized group and the control group received trial by trial feedback in the training phase but the verbalized group also verbalized their knowledge from the training into written instructions. The learning group did not receive any trial by trial feedback in the training phase, instead they were handed the written instructions made by the verbalized group.

The present study supports Enqvist et al.’s (2006) results of higher performance caused by intervention when the verbalized individuals (group 1) demonstrated more accurate judgments overall. According to the learning speed no significant differences were shown even for the individuals that lack training feedback. An explanation could be that no feedback after every trial is necessary if the descriptions handed out by the individuals that verbalized their knowledge as written instructions are structured and clear enough. The control group that received trial by trial feedback but did not have to verbalize their knowledge as written instructions showed lower performance in the training phase than the individuals that verbalized their knowledge. These results strengthen the evidence that verbalization increases learning, which could be explained by the verbal facilitation effects described by Huff and Schwan (2008; 2012). Further, the results clearly demonstrate that it was a great learning advantage to present stimuli analogically rather than propositionally, in contrast to previous studies. Stimuli in the format of analogues were in many aspects more beneficial irrespective of the form of verbalization or absence of feedback, which partly differs from Ainsworth (2006) who argues there is a balanced approach in which both text and pictures contribute in a more or less equal degree to knowledge acquisition.

How the verbalization was formulated in the instructions seems to be important for what kind of knowledge representation is used. This also affects the learning and

use of knowledge representation for the participants that learned from the written verbalization as the verbalization was their only feedback. The participants in this group were affected by the verbalized instructions to store whole exemplars instead of relying on cue abstraction regardless of stimuli presentation. Maybe reading informative stimuli or descriptions of cue abstraction character create an analogue representation in the memory to more easily make a comprehensive whole of the stimuli. The participants could therefore implicitly be forced to make exemplars of the propositions. These exemplars could be remarkably different from the exemplars in the analogue condition but fulfill the aim of an individual reference point. Furthermore, the results of this study show that exemplar-based knowledge could be the regular result of cooperation, even without social interaction.

References:

- Ainsworth, S. E. (2006). DeFT: a conceptual framework for learning with multiple representation. *Learning and Instruction, 16*, 183-198.
- Enqvist, T., Newell, B., Juslin, P., & Olsson, H. (2006). On the role of causal intervention in multiple-cue judgment: Positive and negative effects on learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 163-179.
- Huff, M., & Schwan, S. (2008). Verbalizing events: Overshadowing or facilitation?. *Memory & Cognition, 36*, 392-402.

How do We Judge the Healthiness of Food?

Jacob L. Orquin

*Aarhus University, Department of Business Administration / MAPP,
Denmark*

Contact: jalo@badm.au.dk

The following is a short summary of some thoughts behind the article “A Brunswik lens model of consumer health judgments of packaged foods” appearing in *Journal of Consumer Behaviour*.

In my naiveté I decided I wanted to know how real people judge the healthiness of real products. Not mock-products, but the real deal, the sort of stuff you see in the supermarket. Considering the obesity epidemic in western countries one would think that such a question had already been answered. To my surprise I found that it had not. Despite abundant studies on consumer reactions to health and nutrition information no one had looked at how consumers react to a representative sample of real products. In other words, an enormous body of knowledge had been accumulated on how isolated features of consumer products influence health judgments.

To remedy this I went on a super-size shopping trip and bought all products I could find in the dairy section: cheese, milk, yogurt and butter, a total of almost 200 products. After taking high resolution images of the front and back of each product I began coding each of them on nutrition information, health and nutrition claims, organic claims, brands, product category, flavor and any other product cue I could think of. Equipped with a database of roughly 200 coded product images I collected health judgments from a representative sample of more than 1000 Danish consumers.

To better understand the consumer judgment process and how it relates to the market environment that consumers navigate I opted for a Brunswik lens model. To build the lens model of consumer health judgments I first computed a nutrition score for each product according to WXYfm nutrition model (Scarborough et al., 2007). The nutrition model is used by, for instance, the Australian government when deciding about which products to endorse for public health reasons and thus provides an estimate of each product’s ‘true healthiness’. The main analyses used each of the coded product cues to predict the consumer judgments and nutrition score respectively (see Figure 1).

To my surprise I found that consumers are not all that bad in judging food healthiness, especially when considering that they almost entirely rely on a single cue when forming their judgments: the food category. That is, consumers’ health judgments are determined by whether the product is strawberry yoghurt, butter, full-fat milk etc. and not by any objective nutrition information such as how much fat, protein, or sugar it contains or by whether it bears an official health label. Even more surprising, I found that there was practically no effect of nutrition literacy on accuracy or cue utilization. Consumers with a high level of health and nutrition literacy also rely

almost entirely on the food category. Although one would expect people with more knowledge about health and nutrition to read and use nutrition information, it makes some sense why they would not. It turns out that the food category is an excellent predictor for healthiness and furthermore requires a minimum of effort to process, which cannot be said about nutrition information.

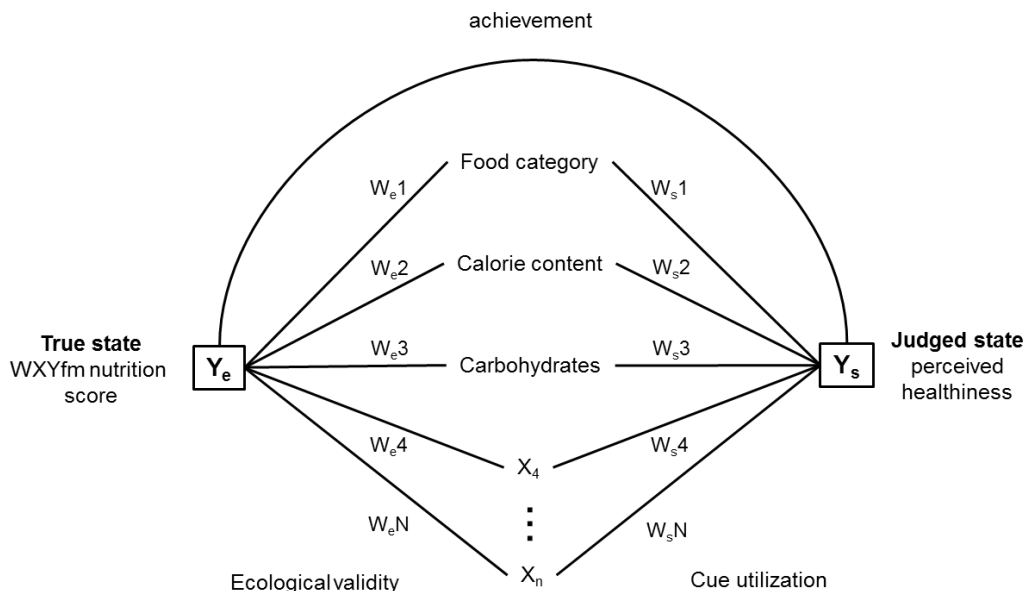


Figure 1. A lens model of consumer judgments of food healthiness.

I believe my study has resulted in a couple of important insights. First, it demonstrates an almost ideal case of *ecological rationality*, i.e. when the structure of bounded rational decisions matches the structure of information in the environment (Todd & Gigerenzer, 2012). Consumers rely on the single most accessible and predictive cue in forming judgments, thus achieving a reasonable level of accuracy with almost no effort. Second, it shows that education has virtually no impact on consumer judgment accuracy or cue utilization. Third, the results suggest that policies aiming at improving health and nutrition literacy may be a waste of resources. The last point is, of course, open to debate and should not serve as an argument for less public education in general. The point being that if policy makers wish to enhance healthy food choice it would probably lead to better results to focus on environmental changes that facilitate replacing one product category with another, e.g., replacing toast bread with whole grain bread. Policies aiming at educating consumers to choose the healthiest product within a category, for instance choosing the healthiest type of toast bread, seem likely to fail.

References:

Orquin, J. L. (2014). A Brunswik lens model of consumer health judgments of packaged foods. *Journal of Consumer Behaviour*, 13(4), 270-281.

Scarborough, P., Boxer, A., Rayner, M., & Stockley, L. (2007). Testing nutrient profile models using data from a survey of nutrition professionals. *Public Health Nutrition*, 10(04), 337-345.

Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. Oxford University Press.

**Do Communication Abilities Affect Promotion Decisions?
Some Data from the C-Suite**

N. Lamar Reinsch, Jr.
Georgetown University, Washington DC,
USA

Jonathan A. Gardner
Home Front Communications, Washington DC,
USA

Contact : ReinschL@georgetown.edu

This paper (published in the January, 2014, issue of the Journal of Business and Technical Communication) was based on an analysis of data from a representative sample of senior executives in U.S. business organizations. The analysis was guided by the Brunswikian lens model.

The survey asked executives (n = 303) who had made a recent decision to promote someone to reflect on the factors that influenced the decision. Responses revealed that in most cases the executive had considered only a few candidates (often having a particular person in mind from the beginning), had considered multiple factors in making the decision, and had found it relatively easy to reach a decision.

In responding to an open ended question about their decision factors, the executives frequently mentioned communication, especially interpersonal communication abilities. Examples included expressions such as “rapport builder,” “ability to work well with people,” and “people skills.”

A more detailed analysis concerned rankings of 23 decision factors collected with the *maximum difference scaling* (MaxDiff) technique. This technique presents a respondent with a series of computer screens and on each screen a request to identify – from a list of five – the most and least important factors. A respondent’s answers are tabulated by computer and what emerges is a complete ranking of all factors.

Results revealed that the most important factors were *past job performance*, and *leadership potential*. Another factor that emerged as extremely important was *strong interpersonal skills* (248 MaxDiff points). Among those factors ranked as average in importance was the factor of *oral communication skills* (109 MaxDiff points). Ranking considerably lower was the factor of *written communication skills* (56 MaxDiff points). These results should be of interest to communication educators and trainers.

We were also interested to see that the age of the respondent affected his or her answers. Older respondents (“Boomers”) rated communication (interpersonal, oral, and written) more highly than did younger respondents (“Gen X”). While this difference could be a maturational effect, we conclude that it is more likely a generational effect, suggesting that younger executives will not change their opinions as they mature.

The Decision Maker as a Planner

Lars Sjö Dahl
*University of Lund, Lund,
Sweden*

Contact: lars.e.sjodahl@gmail.com

Occasionally, decision researchers may frame their problems and scenarios too narrowly (Kahneman & Lovallo, 1993). In medicine, narrow framing of diagnostic questions may be misleading and result in serious diagnostic errors (Zamir, 2014). In order to counteract a similar narrowness in psychological research Brunswik (1956) suggests the following sampling procedure: “representative sampling is extended from the subjects to the objects from individuals to the stimulus situations and tests”(p. 58). Hammond (1948) reminds us that “Brunswik’s main point in connection with sampling is that it should be performed with respect to both subject and object, if we are to generalize in both directions” (p. 530). A similar idea to restrict causal speculations to delimited subject- or situation-areas, called “causal fields”, is suggested by von Wright (1972). In our Brunswikian context, causal fields may also be called “defined ecological domains” from which we sample stimulus situations, problems or observations (Sjö Dahl, 2009, 2011, 2012). However, Brunswik’s representative sampling design has not been applied to any great extent by his fellow researchers. Obviously there are some conditions or circumstances that cause a wider application of Brunswik’s representative sampling to be a bit difficult to follow. We might expect that this inertia to apply a bright idea in a wider context may have something to do with situational circumstances and conditions that restrict the researcher’s work situation. The following aggravating circumstances might very well have something to do with this neglect of Brunswik’s representative design concept:

1. The complexity of the task, the problem to be solved. Difficulties in defining what should be represented on the stimulus side.
2. Environment’s uncertainty over time; ecologies change fairly fast.
3. Interdisciplinary obstacles; a representative sample of situations quite often requires interdisciplinary cooperation and specialists are not always on speaking terms with each other.
4. An astonishing lack of qualitative approaches as complement to quantitative studies.
5. Allocation of research findings to an unintended application level.

Tolman and Brunswik (1935) summarize their article about mental representations and the causal texture of environment as follows: “...this function of local representation has, however, two sub-varieties: 1) objects or situations may function as local representatives of others in that they provide (with the cooperation of the organism) means-objects to the others; these latter to be known as the goals, 2) objects and events may also as local representatives for others in that, being themselves caused by such other objects or events they serve as cues for the latter”

(p. 73). This description of the environment's causal texture is not restricted to an isolated environment. On the contrary, the authors emphasize the interaction between organism and environment with the parenthesis: "(with the cooperation of the organism)". With this broad definition of the concept "representative" it becomes natural to regard our covert behavior, e.g., feelings, emotions, thoughts and decisions as urgent topics for psychological research. However, it is a bit strange that human covert behaviors, i.e. our mental representations, are given so little attention in Brunswikian research. Besides our interesting and successful efforts to study the actual judgment and decision process via pre-defined task descriptions and criteria-defined achievements there is obviously a need for methods to map the individual decision process as a path with a beginning and an end, connected by intermediary mental representations of actions, goals and consequences and feelings, as a chain of mental events, representations and combinations, moving around in time from a start to a goal. Without such a complementing, qualitative approach, Brunswikian research runs the risk of being perceived as a rather mindless enterprise.

Even if our knowledge about the neural correlates to our conscious states will grow rapidly there still remains a need for studying consciousness in terms of subjective content. The neurologist Benjamin Libet has expressed this complementary need as follows: "Our own subjective inner life, including sensory experiences, feelings, thoughts, volitional choices, is what really matters to us human beings" (see Crick, 1994, p. 255). Accepting this qualitative stand means that we see the decision maker as an anticipatory planner who moves mental representations around making currently perceived space change so the initial, surface perception of the task gives way for a deeper space, thereby coming closer to the path's goal, the decision maker's final solution. It is difficult to imagine research about such a path process without special methods adapted to qualitative data recording.

THE INTENTIONAL ASPECT

Human behavior can hardly be described in causal terms unless the goal or intentional aspect is taken into consideration (Tolman & Brunswik, 1935; Miller, Galanter & Pribram, 1960; Eccles, 1970; Dennett, 1997; Greenfield, 2001). This goal aspect could also be described in terms of different planning capacities and processes as suggested by Gulz (1991):

- "A capacity to represent possible situations – in particular, goal and start situations;
- A capacity to represent and evaluate possible events – in particular its own possible actions: to represent their prerequisites and consequences or, in other words, the situations they can transform and how. To do this, the planning system must have some representation of the spatial and causal structure of its environment;
- A capacity to manipulate ideas prior to acting upon them, that is to represent actions and goals without immediately acting upon them" (p. 46) (Gulz' stress).

This author emphasizes that her definition "is intended to be a relatively general definition of planning" (p. 46).

DIFFERENT GOAL-LEVELS

Most, perhaps all, Brunswikian studies present with a predefined goal description, such as evaluating patients on three traits (Nystedt & Magnusson, 1975), or predictions of advertising sales for *Time Magazine* (Ashton, 1982), etc. Qualitative studies as suggested in our text would probably not be restricted to fixed goals, but rather result in a row of proposals, reflecting a wide range of goals, purposes, motives and reasons for alternative actions and decisions. Some of them would be rather abstract or general while others might reflect a concern about concrete, practical matters. There would certainly be a variation of goal levels along the scale from abstract, general to concrete, practical. We would guess that the latter will be in a minority. Goals stated by societies' authorities, like national educational or health agencies, are usually worded at high levels of abstraction. Educational authorities may, for example, state their didactic goals in terms like motivation, pupil activity, concretizing, individualized instruction, and cooperation skills. The teacher, on the other hand, responsible for low-level goals, starts a mental inference chain by asking what the necessary conditions are, the premises for realizing these goals. The relation between different goal levels and inferences to act has been discussed by von Wright (1972) and Anscombe (2002). Obviously, there is a fundamental difference in the method of reasoning between the two different goal levels, the overruling, abstract level and the corresponding executing level. In the former case we speak of theoretical inferences starting with an assertion about how to solve a problem. In the latter case the inference chain is different and we talk about practical inferences. It starts with what is wanted and then goes on to considerations of necessary conditions, called premises, for achieving the goals. This is an inference chain resulting in what Anscombe (2002) calls quasi-imperative inferences, i.e., demands to act, practical inferences or self-administered goals. Anscombe (2002) asks: "Is inference a process? Is infer a psychological verb? Is 'reasoning' a psychological concept? If so, it is perhaps curious that people don't usually put inference and reasoning into lists of mental phenomena. ...But now, how is it that when one considers or examines inferences one has no interest in whether anything like that has gone on in someone's mind. ...It is because we have no such interest that it does come natural to classify inference as a mental content, 'infer' as a psychological verb" (pp. 3-4). It is obvious that Anscombe's discussion about the differences between theoretical and practical inferences is closely connected with different goal-levels' duties and responsibilities. This difference is clearly illustrated in the annual report from the National Swedish Board of Health and Welfare (2011, pp. 73-78). "Individual follow-up plans for the patients, compulsory according to a new Swedish law 2010, were not generally implemented, very rarely were they set up. Among 67 patients' records and patient charts only two such plans were found" (Author's own translation). When considering evaluation results like these one may ask: Can cognitive science contribute to more valid practical inferences, i.e., realizations of premises for better goal achievements? A brief example of how to apply Brunswik's representative design will give us some hints.

Already when deciding about a suitable sample of subjects we have to answer the question: What contextual environment is going to be represented in our stimulus material, our critical incident interview questions? In some way a restricted contextual domain or aspect has to be defined to guide the content of our questionnaire. For our short illustration purpose we select a medical specialty, care of schizophrenia and the

aspect quality of life. A qualitative research approach with a defined domain aspect like quality of life combined with the subjects' comments on questions like those below would give us a surplus of case descriptions, illustrating practical inference chains as well as obstacles for realizing the medical staff's self-administered goals.

1. Why did you react to this situation as being critical when you first became aware of it?
2. How did you act just on the occasion when you became aware of the problem?
3. Could you have acted in some other way just at this moment?
4. Were there other ways to handle this critical situation, good or less suitable? Please motivate your evaluation of options.

ATTITUDES AS OBSTACLES

Sometimes self-administered or quasi-imperative goals may be difficult to realize because of attitude obstacles. As Brunswik's representative design is mainly an idiographic approach, emphasizing the need for taking the individual subject's context, the ecology, into consideration we will briefly mention a similar approach within attitude research. In Ajzen's (1991) article "Accounting for actions in specific contexts. The theory of planned behavior" (p. 181), the focus is on the individual's control and relation to his/her actual environment. The objectively described behavior control does not always correspond to the single individual's intentions and subjectively perceived behavior control (alternatives). The greater this difference between objectively described and subjectively perceived control, the less likely that general, overruling goals will be transformed to low-level, self-administered goals, and put into practice. Without some kind of qualitative complementary studies it would, in such a case, be difficult to transform a high-level goal, for example a law about compulsory follow-up plans for patients, into practical inference chains and premises necessary for goal achievement.

ALLOCATION OF RESEARCH FINDINGS TO UNINTENDED GOAL-LEVELS.

In all science we find a strong ambition to classify or categorize observations in order to summarize and create a common language. Also in everyday life we make extensive use of categories. Why this ability to categorize the world around us? The best answer is probably given by Rosch and Lloyd (1978), who suggest that this capacity has developed as a means of attaining cognitive economy. In other words, categorizing has had survival value in the individual's adaption to the environment. It is very efficient to organize the world around us into restricted numbers of classes or categories, rather than trying to store in our own memory every single, individual example. For further discussion about the use and construction of category systems see Sjö Dahl (2012).

In scientific contexts this preference for generalized knowledge has sometimes been extreme and resulted in neglect of single-case information even though it has been highly relevant for the problem in question. Brunswik's emphasis on situation sampling as well as individual sampling is meant to facilitate causal reasoning about situation-individual interaction. In this sense his representative design is idiographic, and runs the risk of not fitting into broader categories of knowledge. This is a situation

that may result in allocation of research findings to unintended higher goal levels. To clarify: A Brunswikian study (Sjödahl, 1974) covering several years of recorded interviews with nurses at their place of work used a modified critical incident approach to collect case illustrations commented on by the responding subjects. In the interview instruction the ecological domain was clearly described and exemplified as the psychosocial aspect on patient-nurse interaction in surgical and medical hospital departments. Positive and negative episodes were equally often asked for. The reported cases were categorized according to 19 situation categories covering the whole working cycle from hospital admittance to discharge of the patient. Each reported case or script was thus classified by the reporting nurse who also reflected over her own responses and suggested alternative ways to handle the situation. This research report was well appreciated by the intended users of the findings as a document about actual shortcomings in hospital care. However, when it came to practical inferences, i.e., demands to act and premises for realizing local (self-administered goals), the review was extremely negative. Although the low-level goals and practical inferences were always presented in connection with a single event, a reported situation, classified by the subject, for example "this happened during the doctor's round," this situatedness was completely ignored as the following generalized review illustrates: "The nurse should have time for all patients who want to talk with her, she should conduct ego-supporting talks and strengthen their hopes" (Sonesson, 1975, p. 4853) (author's own translation). Nowhere in the original report can this generalization be found. The alternative ways to act are the nurse's own suggestions of how to handle the single event/situation she has just reported. This faulty allocation of qualitative research findings to an unintended, more general goal level may be seen as a displacement heuristic, sweeping Anscombe's (2002) practical inference chains and self-administered local goals under the carpet. Magnusson and Stattin (1998) warn that certain psychological theories about the interaction between individual and environment are often summarized in general, broad, distal categories. Obviously, many forces are at cross purposes with Anscombe's (2002) practical inference chains and Brunswik's idea about representative design and context-dependent research results.

Within some fields like psychiatric diagnostics, the neglect of idiographic, contextual information is alarming. Qualitative data giving meaning to diagnostic labels are often non-existing and patients are routinely labeled with descriptive surface symptoms (Sjödahl, 2011, 2012, 2014).

However, the mutual dependence between psychometric and qualitative research approaches is admirably illustrated in a recent study (Strand, Sjöborg, Stahlmeijer et al., 2013) about medical students' perception of their clinical workplace environment. In their final discussion the authors suggest the following complement to their psychometric analysis: "...additional qualitative data to further deeper understanding of the various elements of optimal and suboptimal workplace learning environment for this group of learners" (p. 1023).

References:

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.
- Anscombe, G. E. M. (2002). Practical inference. In R. Hursthouse, G. Lawrence, & W. Quinn (Eds.), *Virtues and reasons* (pp. 1-34). Oxford: Clarendon Press.
- Ashton, A. H. (1982). An empirical study of budget-related predictions of corporate executives. *Journal of Accounting Research*, 20(2), 440-449.

- Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley: University of California Press.
- Crick, F. (1994). *The astonishing hypothesis. The scientific search for the soul*. New York: Charles Scribner's Sons.
- Dennett, D. C. (1997). *Brainstorms. Philosophical essays on mind and psychology*. London UK: Penguin Books Ltd.
- Eccles, J. C. (1970). *Facing reality*. New York: Springer-Verlag.
- Greenfield, S. (2001). *The private life of the brain: Emotions consciousness, and the secret of the self*. London: Penguin Books Ltd.
- Gulz, A. (1991). *The planning of action as a cognitive and biological phenomenon*. Lund, Sweden: Lund University Cognitive Studies, ISBN 91-628-0224-0.
- Hammond, K. R. (1948). Subject and object sampling: A note. *Psychological Bulletin*, 45, 530-533.
- Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39(1), 17-31.
- Magnusson, D., & Stattin, H. (1998). Person context interaction theories. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology. Vol 1: Theoretical models of human development* (pp. 685-759). New York: Wiley.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Henry Holt.
- National Swedish Board of Health and Welfare (2011). *Annual Evaluation Report*, 73-78. (In Swedish).
- Nystedt, L., & Magnusson, D. (1975). Integration of information in a clinical judgment task, an empirical comparison of six models. *Perceptual and Motor Skills*, 40(2), 343-356.
- Rosch, E., & Lloyd, B. B. (1978). *Cognition and categorisation*. Hillsdale, NJ: Erlbaum.
- Sjödahl, L. (1974). *Att vårda människor (Care of people)*. Stockholm: Natur och Kultur. (In Swedish).
- Sjödahl, L. (2009). Brunswik's lens model adapted to aspect representation. *The Brunswik Society Newsletter*, 24, 37-44.
- Sjödahl, L. (2011). Brunswik's relational, environment-organism approach applied to schizophrenic symptoms: An etiological model. *The Brunswik Society Newsletter*, 26, 40-44.
- Sjödahl, L. (2012). *Psychiatric diagnoses – are they to be trusted? First listen to the patient!* Visby, Sweden: Nomen Publishing.
- Sjödahl, L. (2014). Path dependence and routines: A threat to capability development. *Diagnosis*, 2, 183-188.
- Sonesson, B. (1975) Värdefullt dokument om patientvård. (A valuable document on patient care). *Läkartidning* 49, 4852-4853. (In Swedish).
- Strand, P., Sjöborg, K., Stalmeijer, R., & al. (2013). Development and psychometric evaluation of the Undergraduate Clinical Education Environment Measure (UCEEM). *Medical Teacher*, 35, 1014-1026.
- Tolman, E. D., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review*, 42, 43-77.
- von Wright, G. H. (1972). On so-called practical inference. *Acta Sociologica* 15(1), 39-53.
- Zamir, E. (2014). The bias of the question posed: a diagnostic "invisible gorilla". *Diagnosis*, 1(3), 245-248.

**The Structure of Brunswik's "Lens":
A Media-Historical Analysis**

Martin Wieser*Sigmund Freud Privat Universität Berlin,
GER**Contact: martin.wieser@sfu-berlin.de*

Since the 1980s, images and visual media have become an area of major interest in the history and theory of science, science and technology studies and sociology of science. Images, diagrams and iconic material are used widely in the production and dissemination of scientific knowledge both within and outside the laboratory. In contrast to a long-standing tradition of understanding the "core" of scientific knowledge as a matter of written sentences, which still dominates contemporary history of psychology, the new historiography acknowledges visual media as distinct parts of scientific knowledge in their own right (cf. Mitchell, 1995; Daston & Galison, 2010).

With an academic background both in philosophy of science and psychology, my research interest focuses on the development of a media-centered historiography of psychology. Two earlier publications were concerned with philosophical and epistemological implications that go along with psychology's increasing use of visual media (Wieser & Slunecko, 2013a, 2013b). My first historical case study on the roots of psychological imagery was devoted to Freud's early visualization practice, which he was trained in at the Viennese Institute of Physiology under Ernst Brücke, and its profound influence on Freud's later image of the "psychical apparatus" (Wieser, 2013).

As he was a long-time student and assistant at the Viennese Department of Psychology, the work of Egon Brunswik seemed an obvious object of inquiry for the advancement of an image-oriented historiography of psychology. Besides his contributions on statistics, methodology and probabilism, his image of the "lens" probably represents his most important legacy within experimental psychology, a theory of human perception which fundamentally builds upon geometrical methods to visualize the processes of human perception and cognition. What struck me most was the apparent contrast between Brunswik's symmetrical and clear-cut visualization of the "lens" in his habilitation thesis from 1934 (Brunswik, 1934) and Heider's sketch of the "lens" in his private notebook (Heider, n.d./1987, Vol. 2, p. 149). Furthermore, Brunswik's own presentation changed profoundly after his migration (compare Brunswik, 1952, p. 20). While Heider's sketch of the "lens" visualized the relation of the "phenomenon" and "thing" and, most prominently, the layer of the "medium" which connects both of them, Brunswik replaced the concept of "phenomenon" with "image" ("Abbild") and used a geometrical, symmetrical outline for his outline. Over the next decades, the "subjective" side of perception began to vanish in Brunswik's model. In its place, we soon find a "perception-reaction" (Tolman & Brunswik, 1935), a "perceptual response" (Brunswik, 1940) and, in its last version, a series of "initial" and

“terminal variables” which make up a completely objective “functional unit of behavior” (Brunswik, 1952).

The difference between these images appeared even more puzzling as Heider (and sometimes his teacher Meinong) is frequently cited by Brunswik and his followers as the main inspiration for Brunswik’s conception of the “lens”, although Heider harshly criticized Brunswik’s methodology as “stuck in the early views of the ‘Vienna Circle’” (Heider, 1983, p. 132) and argued that “Brunswik does not (curiously!) consider what is necessary to get from atomistic proximal stimulus to the distal world - this step demands more than a mathematical putting together of single items” (Heider, n.d./1987, p. 166).

To gain a better understanding of why Brunswik changed the outline, structure and key components of Heider’s “lens,” I drew upon Frederic Bartlett’s book *Remembering* (Bartlett, (1932/1995) and his understanding of visual transformations as an effect of the application of different “schemes” which imply different sets of norms, beliefs and presuppositions and which are acquired through and embedded within different social collectives.

To sum up my arguments briefly (which are presented in detail, along with the sketches of the lens, in Wieser, 2014): Heider and Brunswik belonged to very different groups of academic psychologists. Heider represented a type of textbook psychologist who was more concerned with conceptual and epistemological questions than the execution and interpretation of singular psychological experiments. He did not ground his arguments on empirical data, as he was concerned with the conceptual foundation and presuppositions on which these data were gathered. His research was closer to phenomenology, essayistic in style and less bound to a certain methodology.

Brunswik’s work, on the other hand, was devoted to the advancement of experimental methods and the proper statistical interpretation of numerical data which were acquired through his experiments. In contrast to Heider, Brunswik was very close to the Vienna Circle and the movement for the “unity of science”. After his migration to Berkeley, Brunswik had to adapt to the cognitive “schemes” which dominated the research of his new social peers, behaviorism and logical positivism, to gain acceptance within his new community. The physicalist ideals of quantification, objectification, operationalization and non-introspectionist experimentation were very strong within these academic collectives, while “metaphysical” problems were supposed to be eliminated from scientific investigations. The various versions of Brunswik’s “lens” are documents of these changing social contexts. Brunswik transformed Heider’s sketchy drawing into a clear-cut geometrical diagram, and replaced philosophical and media-theoretical questions from Heider’s “Thing and Medium” (1926/1959) with experimental tasks for his psychology of “objective relations” (Brunswik, 1937).

In the discussion of his experiments on “serial reproduction,” Bartlett argued that human perception and cognition are structured by sets of beliefs, norms and presuppositions which are not invented by a sole individual but products of a social collective. When images or stories circulate between groups and collectives, these “schemes” exert their influence by transforming strange and “meaningless” cultural elements until they become “meaningful.” In the case of the “lens” a very similar process occurred: Heider’s “lens” addressed an audience of philosopher-psychologists

which were open to his epistemological and phenomenological perspectives, while Brunswik had to eliminate all “subjective” or phenomenological facets of the “lens” to make it compatible with the norms of logical positivism and neobehaviorism. None of the various changes and transformations of the “lens” was based on new empirical findings or experimental data, as each one of them represents a normative guide on how empirical investigations should be performed and interpreted, not purely descriptive accounts of human perception and cognition.

Altogether, my media-historical account was designed to uncover and analyze the social and epistemological conditions that shaped the manifold versions of the “lens.” Its remarkable history documents the struggle of academic psychology with philosophy, epistemology, phenomenology and the search for a unified and exact methodology during the first half of the 20th century. The intellectual origin and later migration of Heider and Brunswik, the decline of phenomenology, the rise of logical positivism, the dominance of neobehaviorism and the dawning of cybernetics within academic psychology all left their marks in the images of the “lens,” whose history and meaning cannot be properly understood without the various social contexts that created and transformed it (p. 1023).

References:

- Bartlett, F. C. (1995). *Remembering: A study in social and experimental psychology*. Cambridge, United Kingdom: Cambridge University Press. (Original work published 1932)
- Brunswik, E. (1934). *Wahrnehmung und Gegenstandswelt*. [Perception and the world of objects]. Vienna, Austria: Deuticke.
- Brunswik, E. (1937). Psychology as a science of objective relations. *Philosophy of Science*, 4, 227-260.
- Brunswik, E. (1940). Thing constancy as measured by correlation coefficients. *Psychological Review*, 47, 69-78.
- Brunswik, E. (1952). *The conceptual framework of psychology*. *International encyclopedia of unified science*. Chicago, IL: University of Chicago Press.
- Daston, L., & Galison, P. (2010). *Objectivity*. New York, NY: Zone.
- Heider, F. (1959). Thing and medium. In G. Klein (Ed.), *On perception, event structure, and psychological environment. Selected papers by Fritz Heider* (pp. 1-34). New York, NY: International Universities Press. (Original work published in German 1926)
- Heider, F. (1983). *The life of a psychologist: An autobiography*. Lawrence, KS: Kansas University Press.
- Heider, F. (1987). *The notebooks* (Vol. 2). München, Germany: Psychologie Verlags Union.
- Mitchell, W. J. T. (1995). *Picture theory: Essays on verbal and visual representation*. Chicago, IL: University of Chicago Press.
- Tolman, E. C., & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review*, 42, 43-77.
- Wieser, M. (2013). From the eel to the ego. Psychoanalysis and the remnants of Freud's early scientific practice. *Journal of the History of the Behavioral Sciences*, 49(3), 259-280.
- Wieser, M. (2014). Remembering the “lens.” Visual transformations of a concept from Heider to Brunswik. *History of Psychology*, 17(2), 83-104.
- Wieser, M., & Slunecko, T. (2013a). Images of the invisible. An account of iconic media in the history of psychology. *Theory & Psychology*, 23(4), 435-457.
- Wieser, M., & Slunecko, T. (2013b). (Theories are) More than words. Why images are significant to theoretical psychology. In A. Marvakis, J. Motzkau, D. Painter, R. Ruto-Korir, G. Sullivan, S. Triliva, & M. Wieser (Eds.), *Doing Psychology under new conditions* (pp. 209-306). Concord, ON: Captus Press.

News from Nebraska

Bob Wigton & Tom Tape

*University of Nebraska Medical Center, Omaha, Nebraska,
USA*

Contact: wigton@unmc.edu, tgtape@unmc.edu

Along with Esther Kaufmann, we presented a short course at the meeting of the European Society of Medical Decision Making in Antwerp, Belgium this past spring. The course was called "The Psychology of Medical Decision Making: How to Analyze Clinical Judgment Using a Lens Model Approach." It was an intermediate level course covering the history of the lens model and how to construct lens model studies. We greatly enjoyed the meeting and took some extra time to visit some of the great art exhibits in Belgium (plus a trip to Amsterdam). The PowerPoint of the short course is available on request (please sent the request to Bob or Tom).

Over the past several years, we have been working with a research group of emergency medicine physicians in Wisconsin to study the factors that affect whether they prescribe antibiotics for acute respiratory illness. We found that their weighting of clinical factors on paper cases was nearly identical to a group of primary care physicians we had studied previously in Colorado (there is a significant problem with overuse of antibiotics in acute respiratory infection, particularly for acute bronchitis which is mostly viral and does not benefit from antibiotics). Additional information you will find at: Safdar, N. S., Tape, T. G., Fox, B., Svenson, J. E., Wigton, R. S. (2014). Factors Affecting Antibiotic Prescribing for Acute Respiratory Infection by Emergency Physicians. *Health, 6*, 774-780, <http://dx.doi.org/10.4236/health.2014.68099>.

We are also working with this group to examine the importance of barriers to use of currently available methods for preventing hospital transmission of *clostridium difficile* infection.

**The German Research Program KoKoHs –
Theoretical Concept, Assessment Design, and Validation Approach
in Modeling and Measuring Competencies in Higher Education**

Olga Zlatkin-Troitschanskaia, Christiane Kuhn, & Miriam Toepper
Johannes Gutenberg University Mainz,
GER

Contact: Istroitschanskaia@uni-mainz.de

The Research Program KoKoHs. The research program KoKoHs is a German research initiative on “Modeling and Measuring Competencies in Higher Education”, funded by the Federal Ministry of Education and Research. It aims to provide systematic, internationally compatible and visible fundamental research on competency assessment and development in higher education in Germany (see http://www.kompetenzen-im-hochschulsektor.de/index_ENG.php; Zlatkin-Troitschanskaia, Kuhn & Toepper, in press). The first funding phase is running from 2011–2015. The program includes 24 cross-university project alliances encompassing nearly 70 single projects with approximately 220 researchers at more than 50 institutions of higher education in 14 German federal states as well as in Austria.

Earlier approaches to competency assessment in Germany were mostly limited to prerequisites for higher education admissions tests, data on learning opportunities, and subjective measures (see Kuhn & Zlatkin-Troitschanskaia, 2011). The KoKoHs research program aims to systematically model and assess domain-specific and generic competencies in higher education. KoKoHs projects take into account curricular and job-related requirements, transform theoretical competency models into suitable measuring instruments, and validate test score interpretations. KoKoHs focuses on selected domains of study, including economic and social sciences, engineering sciences, educational sciences, and teacher training in STEM subjects (science, technology, engineering, and mathematics).

During the first funding phase, KoKoHs projects defined competencies holistically as latent cognitive and affective-motivational underpinnings of performance (see Ewell, 2005; Rycher, 2004; Winch & Foreman-Peck, 2004; Weinert, 2001). Models of cognitive abilities and skills were operationalized through measuring instruments and tested in empirical assessments. Validation efforts aimed to establish validity of interpretation of the evidence and focused on the key question: What can we infer from the cognitive representations elicited by the assessment about the actual competencies of individual students? This approach always includes a challenge: The underlying abilities and skills – ideally also the corresponding affective-motivational aspects – need to be operationalized through representative, practice-oriented, and often domain-specific tasks; assessments need to represent specific situational contexts and be free of potential bias, such as measurement errors or influences of construct-irrelevant test-taking behavior.

In the following, we outline key aspects of the KoKoHs program, including the concept of competency, the assessment design, important validity issues, and future perspectives.

Concept of Competency

Following Weinert's holistic definition (2001), the KoKoHs program generally defined competencies as "cognitive abilities and skills that individuals possess or acquire in order to solve certain problems as well as the aligned motivational, volitional and social dispositions and skills to apply the solutions in different situations successfully and responsibly" (pp. 27-28), specifically competencies acquired in higher education. For practical reasons, KoKoHs projects focused on cognitive abilities and skills and specified them for their respective study domains. According to Klieme et al. (2008), competency can be narrowed down to cognitive dispositions that are context-specific, learnable, and related to practical situations. The aligned non-cognitive aspects, such as motivational, volitional, and social dispositions, were assessed separately.

Competency constructs were further defined with the following characteristics: dispositional (not performance), latent, relatively stable over time and within situations, influenced by dynamic state components, changeable in level through learning and forgetting, domain-specific, situation-dependent, comprising multiple cognitive and non-cognitive abilities and skills, and evidence-based.

Assessment Design and Validation Approach

The general assessment framework in KoKoHs was based on the Assessment Triangle by Pellegrino, Chudowsky, and Glaser (2001, p. 44), which covers three interdependent aspects that are fundamental to many assessments: "a model of student cognition and learning in the domain, a set of beliefs about the kinds of observations that will provide evidence of students' competencies, and an interpretation process for making sense of the evidence". These three aspects corresponded with key objectives of KoKoHs:

1. defining the construct to be assessed (cognition),
2. developing and using suitable models and measuring instruments (observation), and
3. drawing valid inferences from the assessment data (interpretation).

The Assessment Triangle provided the cornerstones for an assessment connecting theoretical constructs of students' individual competencies with empirical evidence, that is, estimates based on limited instances of students' abilities and skills, in an argument-based approach of "reasoning from evidence" (Mislevy, 1994). For more specific, practical orientation, KoKoHs projects adopted the evidence-centered assessment approach and test development concept (Mislevy & Haertel, 2006; Hattie, Bond & Jaeger, 1999), which includes the following steps (see Brückner, Zlatkin-Troitschanskaia & Förster, in press):

- *Domain analysis and domain modeling:* In the assessment of competencies in higher education, important initial steps included

analyzing and defining the domain and modeling the domain-specific construct to be assessed.

- *Assessment framework*: Next, an assessment framework was defined, which served for operationalizing the theoretical model and developing items for the test instruments.
- *Assessment implementation*: The test instruments were tested empirically, and item responses were converted into test scores, which were aggregated.
- *Assessment delivery*: The test scores were analyzed using various psychometric models. Analyses always included evaluations of fit of the data to the theoretical constructs and to the corresponding test score interpretations. The conclusive evaluation of the test instruments with regard to various validation criteria is under way and will serve as a basis for further decisions.

Validation is of key importance in KoKoHs. Validation efforts follow the International Standards for Educational and Psychological Testing (AERA, APA & NCME, 2004; 2014). These provide a joint methodological focus for very heterogeneous projects, while still leaving enough room for individual projects' aims and needs. So far, most KoKoHs projects have focused on the validation criteria of "test content" and "internal structure"; some have also analyzed "response processes" and "relations to other variables" such as courses attended at university (see Brückner et al., in press).

Future Perspectives

So far, the KoKoHs program has addressed methodological challenges such as systematically designing or adapting items, considering framework conditions, such as time, method, and format, analyzing data with complex psychometric methods, confirming psychometric quality criteria, and undertaking comprehensive validation. For the next funding phase of the KoKoHs program, the models of competency structures, the assessment designs, and the measuring instruments developed and tested so far will provide a solid basis for in-depth longitudinal multilevel analyses in field experimental validation studies.

Researchers familiar with the Brunswikian perspective will recognize that the orientation of the KoKoHs program shares important characteristics with Brunswik's framework, specifically the importance of domain analysis and domain modeling and the notion that assessments need to represent specific situational contexts. We would welcome discussion and collaboration regarding the possibility of the application of a lens model approach or another Brunswikian method or concept to the domain of educational assessment.

References:

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (AERA, APA & NCME). (2004, 2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Brückner, S., Zlatkin-Troitschanskaia, O., & Förster, M. (in press). Relevance of test adaptation and validation for international comparative research on competencies in higher education – A methodological overview and example from an international comparative project within the KoKoHs

- research program In F. Musekamp & G. Spöttl (2015) (Eds.), *Competence in higher education and the working environment. National and international approaches for assessing engineering competence*. Frankfurt am Main, Bern, Bruxelles, New York, Oxford, Warszawa, Wien: Peter Lang (Vocational Education and Training: Research and Practice).
- Ewell, P. T. (2005). Can assessment serve accountability? It depends on the question. In J. C. Burke & Associates (Eds.), *Achieving accountability in higher education* (pp. 1-24). San Francisco, CA: Jossey-Bass.
- Hattie, J., Jaeger, R. M., & Bond, L. (1999). Persistent methodological questions in educational testing. *Review of Educational Research, 24*, 393-446.
- Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In E. Klieme, J. Hartig & D. Leutner (Eds.), *Assessment of competences in educational contexts*. Washington: Hogrefe and Huber.
- Kuhn, C., & Zlatkin-Troitschanskaia, O. (2011). *Assessment of competencies among university students and graduates – Analyzing the state of research and perspectives*. Johannes Gutenberg University Mainz: Arbeitspapiere Wirtschaftspädagogik [working paper: business education], 59.
- Mislevy, R. J. (1994). *Test theory reconceived: CSE technical report 376*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Mislevy, R. J., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement: Issues and Practice, 25*(4), 6-20.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.) (2001). *Knowing what students know: The science and design of educational assessment*. Washington, D. C.: The National Academies Press.
- Rychen, D. S. (2004). Key competencies for all: An overarching conceptual frame of reference. In D. S. Rychen, & A. Tiana (Eds.), *Developing key competencies in education: Some lessons from international and national experience* (pp. 5-34). Paris: UNESCO.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies*. Seattle, WA: Hogrefe and Huber Publishers.
- Winch, C., & Foreman-Peck, L. (2004). Editorial. *Policy Futures in Education, 2*(1), 1-4.
- Zlatkin-Troitschanskaia, O., Kuhn, C., & Toepper, M. (in press). Modelling and assessing higher education learning outcomes in Germany. In H. Coates (2014) (Ed.), *Assessing learning outcomes: Perspectives for quality improvement*. Frankfurt/Main: Peter Lang Publishing Group.

To get information about future Newsletters you are welcome to join (free of charge) the email list of the Brunswik Society at:

<http://www.brunswik.org/>